

Rositsa Dekova

Lexical Encoding of Verbs in English and Bulgarian

Thesis for the degree doctor artium

Trondheim, December 2006

Norwegian University of Science and Technology
Faculty of Arts
Department of Modern Foreign Languages

NTNU

Norwegian University of Science and Technology

Thesis for the degree doctor artium

Faculty of Arts

Department of Modern Foreign Languages

© Rositsa Dekova

ISBN 82-471-8233-5 (printed version)

ISBN 82-471-8232-7 (electronic version)

ISSN 1503-8181

Doctoral theses at NTNU, 2006:226

Printed by NTNU-trykk

For Natalia and Anton

Acknowledgments

First and foremost, I am indebted to my supervisor prof. Mila Dimitrova-Vulchanova for always being patient while wisely guiding the way and nurturing the scholar in me.

I would like to acknowledge the role of my high school teachers Mrs Yaneva, Mrs Gočeva, and Mrs Vŭlkanova, who gave me a glimpse of the world of linguistics, and Penka Stateva and Iliyana Krapova, my university lecturers, who opened the door to that world for me and saw I belong there.

I want to thank prof. Arne Halvorsen and Bjørn Kolstad for their help and concern in administering my stay in the university. My thanks go also to my colleagues from the Department of Modern Languages and the Linguistic Department at NTNU and the Department of Computational Linguistics at the Bulgarian Academy of Sciences, who were my friends as much as my dear colleagues. Special thanks to Anja and Zara for sharing not only their office space but also their experience and good will. Thanks to Tsvetana and Petter for always being there for me and to all my friends for their support and encouragement.

Thanks again to all the people who participated in the preliminary tests and in the online studies, and to Liliana, who helped with the arrangements of the online version of the experiments.

*This work has benefited from discussions at the Ninth Nordic Conference on English Studies, 27-29 May, 2004, Aarhus, Denmark; the symposium *The Lexicon: its status in the theory of language*, Åbo Akademi University, Turku, 18-19 November, 2004; the 3rd International Workshop on generative Approaches to the Lexicon, 19-21 May, 2005, Genève, Switzerland; and FDSL6, 30.11– 2.12, 2005, Potsdam, Germany, where I have presented different stages of this project. I would like to thank the participants in these discussions for their valuable comments.*

I want to thank Annjo and Rurik Greenall for being my "Norwegian" family and always taking good care of me.

I would like to express my gratefulness to my parents and to my parents-in-law for their unceasing support without which I could neither start nor finish this project.

Last but not least, to my husband and daughter, who always believed in me: Thank you for all you sacrificed for me!

Contents

1	Introduction	1
1.1	A Shortcut to Bulgarian Syntax	2
<i>1.1.1</i>	<i>Word Order in Bulgarian</i>	4
<i>1.1.2</i>	<i>The subject – its presence or absence in the Bulgarian sentence</i>	5
<i>1.1.3</i>	<i>Direct Object in Bulgarian</i>	7
<i>1.1.4</i>	<i>Indirect Object in Bulgarian</i>	10
1.2	Brief Typological Juxtaposition of English and Bulgarian Verbs	14
<i>1.2.1</i>	<i>Bulgarian verbal morphology vs. English verb particles</i>	14
<i>1.2.2</i>	<i>Causation and Causative Alternations in English and Bulgarian</i>	15
<i>1.2.3</i>	<i>Se-constructions in Bulgarian</i>	18
1.3	Conclusions	20
2	Current Relevant approaches to the Syntax-Semantics Interface	21
2.1	Verb Classes and Alternations	22
2.2	Construction Grammar	25
<i>2.2.1</i>	<i>More on the Resultative construction</i>	29
<i>2.2.2</i>	<i>The Way-construction – Constructional Polysemy?</i>	33
2.2.2.1	The Means Interpretation	34
2.2.2.2	The Manner Interpretation	35
2.2.2.3	The etymology of the way construction – the diachronic evidence	36
2.3	Generative Lexicon	39
<i>2.3.1</i>	<i>The Lexicon as a multilevel system of semantic representations</i>	40
2.3.1.1	Argument Structure	40
2.3.1.2	Event Structure	42
2.3.1.3	Qualia Structure	43
2.3.1.4	The Lexical Inheritance Structure	45
<i>2.3.2</i>	<i>The generative semantic operations</i>	45

2.3.2.1	TYPE COERCION	45
2.3.2.2	SELECTIVE BINDING	47
2.3.2.3	CO-COMPOSITION	49
2.4	Conceptual Semantics and the Multi-tiered Representation of CS	50
2.4.1	<i>Conceptual Semantics</i>	51
2.4.2	<i>Conceptual Structure and the lexical items</i>	53
2.4.3	<i>The tiered representation of Conceptual Structure (CS)</i>	54
2.4.3.1	The thematic tier	56
2.4.3.2	The action tier	58
2.4.3.3	The temporal tier	59
2.5	Conclusions	62
3	Theoretical Background	64
3.1	Lexical Encoding of Participant Information	65
3.1.1	<i>Lexically encoded vs. non-lexically encoded participant information</i>	65
3.1.2	<i>Possible syntactic criteria and their inadequateness</i>	67
3.1.2.1	Syntactic obligatoriness	67
3.1.2.2	Iteration of participants	69
3.1.2.3	Head-dependence and head selectivity	71
3.1.3	<i>Semantic criteria</i>	72
3.1.3.1	Semantic obligatoriness	73
3.1.3.2	Verb class specificity	73
3.1.4	<i>The Semantic criteria at Test</i>	74
3.1.4.1	Tests targeting the semantic obligatoriness condition	74
3.1.4.2	Tests targeting the semantic selectivity condition	76
3.2	Verb Sense and Subcategorization Preferences	78
3.2.1	<i>Verb sense</i>	78
3.2.2	<i>Subcategorization biases</i>	80
3.2.3	<i>Corpus analyses</i>	80
3.2.4	<i>Empirical Data</i>	81
3.2.4.1	Offline norming experiments	81
3.2.4.2	Online reading experiment	83

3.3 The Sign Model and the Lexical Encoding of Verbs	86
3.3.1 <i>The structure of the Sign</i>	87
3.3.2 <i>The Cell</i>	88
3.3.2.1 Aspectual Specification	89
3.3.2.2 Element Specification	90
3.3.3 <i>The Dimensions</i>	90
3.3.3.1 Force	91
3.3.3.2 Monodevelopment	92
3.3.3.3 Conditioning	93
3.3.3.4 Control	94
3.3.3.5 Criteriality	97
3.3.4 <i>Realization of Criterial Elements</i>	100
3.3.4.1 Canonical positions. Canonically retrievable positions	100
3.3.4.2 Suppression of elements and the survival of the criterial	102
3.3.4.3 Conceptually present participants	103
3.4 Conclusions	103
4 Corpus Data Analyses	105
4.1 Methodology of the analyses and terminology used	106
4.1.1 <i>Methodology</i>	107
4.1.2 <i>Terminology</i>	109
4.1.2.1 Source or Initiator	110
4.1.2.2 Source Extension	111
4.1.2.3 Absorber or Limit	113
4.1.2.4 Place of Contact	114
4.1.2.5 Path values	117
4.1.2.6 Other terms used in the analyses	119
4.2 Verb Grouping	123
4.2.1 <i>Verbs of Contact</i>	123
4.2.2 <i>Verbs denoting a Conditioned event</i>	131
4.2.2.1 Verbs of sustained contact	132

4.2.2.2	Verbs of Disintegration	136
4.2.2.3	Verbs of Initial contact and Ejection	141
4.2.2.4	The Dual Lexicalization Pattern	148
4.3	Extended uses of the examined verbs	154
4.4	Conclusions	159
5	Online Sentence Continuation Studies	160
5.1	Design and Methodology	161
<i>5.1.1</i>	<i>Participants</i>	161
<i>5.1.2</i>	<i>Stimuli</i>	162
<i>5.1.3</i>	<i>Procedure</i>	165
5.2	Analyses of the Results	166
5.3	Results and Discussion	168
<i>5.3.1</i>	<i>Results for Verbs of Contact</i>	169
5.3.1.1	Results for stimuli of the type Subject _[Initiator] Verb ____	169
5.3.1.2	Results for stimuli of the type Subject _[Initiator] Verb Object ____	174
5.3.1.3	Results for stimuli of the type Subject Verb ____	177
<i>5.3.2</i>	<i>Results for Verbs Denoting a Conditioned Event</i>	179
5.3.2.1	Results for stimuli of the type Subject _[Initiator] Verb ____	179
5.3.2.2	Results for stimuli of the type Subject _[Initiator] Verb Object ____	184
5.3.2.3	Results for stimuli of the type Subject Verb ____	188
5.4	Conclusions	193
6	Concluding Remarks	195
	Bibliography	199
	Table of Transliterations	208
	Appendix A	209
	Appendix B	210
	Appendix C	213
	Appendix D	216
	Appendix E	219

1. Introduction

This project investigates the mapping of elements from the semantic representation of verbs onto overt syntactic realization and how this information might be lexically encoded cross-linguistically. The research explores various phenomena arising on the interface of conceptual structure with syntax. The analyses are based on empirical data and experimental evidence from two Indo-European languages – English and Bulgarian. English was selected since it provides with a number of intensively discussed topics on the syntax-semantics interface, thus offering a venue for comparison and juxtaposition with Bulgarian for which the literature is not so comprehensive and works by Dimitrova-Vulchanova (1996/99, in press), Kasabov (1990), and Koeva (1998, 2005) are among the few exceptions.

For the purposes of the project I selected a set of verbs¹ representing several basic verb types in English and Bulgarian and I examined their semantic properties with an account of their syntactic distribution. Special attention was paid to subgroups of the so-called class of *Verbs of Contact by Impact* (as defined in Levin, 1993) and some verbs which include motion (in Levin's classification they fall in the group of *Throw Verbs*). These verbs were of particular interest because of the diversity of patterns of alternation that they allow, as well as the constraints they pose on their syntactic environments.

¹ An overview of the English and the Bulgarian verbs discussed in this work is available in Appendix A.

Currently, it is widely accepted across the different linguistic theories² that the meaning of a verb is closely related to the verb's morpho-syntactic realization. Thus, it is generally assumed that the semantics of verbs reveals to a great extent the possible syntactic patterns that they can display, and in reverse, the syntactic behaviour of a verb gives us major clues towards the information that may be lexically encoded in it. Therefore, analyses of corpus data and the results from different online linguistic studies were used as reliable sources of finding the information that is encoded in verbs, as well as testing how this information is utilized by native speakers in online language production. This approach is in line with several recent theoretical assumptions, as well as with some experimental studies discussed in Chapter 3.

To set a venue and prepare the reader for the discussion of the empirical data analysed in this work, this introductory chapter brings forward some of the distinctive features and characteristic properties of Bulgarian (section 1.1) with respect to the syntactic realization of the various semantic participants to be discussed in Chapters 4 and 5. In addition, a concise juxtaposition of English and Bulgarian verbs is included to outline some of the typological differences and similarities between the two languages.

1.1 A Shortcut to Bulgarian Syntax

Bulgarian is a member of the South-Slavic language group, together with Macedonian, Serbo-Croatian, and Slovenian. Thus, among some of the most prominent characteristic features of Bulgarian are its SVO word order, its classification as a pro-drop language, its highly inflected verb system (including also aspectual specification among other factors), and remnants from once well developed case system, currently discernible only in the different pronominal forms for the Nominative, the Accusative, and the Dative cases.³

² Some of the recent theoretical approaches which take this agenda are discussed in the next two chapters of this work.

³ *Case* is a highly disputed category for Bulgarian. Nevertheless, it is generally considered a morphological category defining the relations between a nominal and another nominal or an activity (cf. Kucarov, 1997 for discussion.)

In addition, Bulgarian is also one of the core members in the Balkan Sprachbund, which embraces the languages spoken on the Balkan Peninsula⁴ like Romanian, Albanian, Serbian, Macedonian, and Greek, together with some minority languages like Arli Romany, Aromanian, Megleno-Romanian and Istro-Romanian which are considered in the centre of the Balkan language continuum (cf. Assenova 1989/2002; Dimitrova-Vulchanova & Vulchanov (forthcoming); Tomić 2004).

As a Balkan language, Bulgarian displays all the six syntactic properties, widely acknowledged to be the most prominent distinctive features of the Balkan Sprachbund. These are discussed in Assenova (1989/2002) and Dimitrova-Vulchanova & Vulchanov (forthcoming) among others, and repeated in (1) below as listed in Tomić (2004).

- (1) a. Substitution of synthetic declension markers by analytic ones;
- b. Grammaticalization of the category of definiteness through postpositive definite articles;
- c. Pronominal doubling of objects;
- d. Analytic expression of futurity;
- e. Analytic perfect with a *have*-auxiliary;
- f. Loss of the infinitive and its substitution by subjunctive clauses.

All these features can be identified in the Bulgarian examples to be discussed in the following chapters. In this introductory chapter, however, only those properties will be presented that have direct relation to the syntactic realization of the semantic participants encoded in the lexical representation of verbs.

⁴ However, not all the languages on the Balkan Peninsula are subsumed in this union or exhibit the same degree of affiliation to the Balkan Sprachbund. Thus, Turkish was never properly included in the Union (cf. Dimitrova-Vulchanova & Vulchanov (forthcoming) and Assenova (1989/2002) for discussion on the cluster of properties displayed by the members of the Balkan Sprachbund). Alternatively, only dialects and not the official languages could be considered (cf. Tomić (2004) for overview and discussion of the different approaches).

1.1.1 Word Order in Bulgarian

Like all the languages from the Slavic group and many other Indo-European languages (including the Germanic languages as English), Bulgarian is generally believed to exhibit an SVO word order, as illustrated in the example in (2) below.

- (2) a. Toj ŝte vidi lodkata.⁵
 he will see boat-the
 a' He will see the boat.

However, the word order displayed in many of the Bulgarian sentences, is relatively free, yet dependent on information structure properties, as illustrated in the examples in (3) below, which are syntactic variations of the sentence in (2).

- (3) a. Toj ŝte ja vidi (lodkata).
 he will her-cl. see (boat-the)
 b. Ŝte ja vidi (toj) (lodkata) *(Right Dislocation)*
 will her-cl. see (he) (boat-the)
 c. Ŝte ja vidi (lodkata) (toj) *(Subject Dislocation)*
 will her-cl. see (boat-the) (he)
 b. Lodkata toj ŝte (ja) vidi. *(Left Dislocation)*
 boat-the he will (her-cl.) see
 c. Lodkata ŝte (ja) vidi toj. *(Clitic Left Dislocation)*
 boat-the will (her-cl.) see he

In addition, there are certain constraints on the Bulgarian word order with respect to the position of the pronominal and the verbal clitics in the linearization of the sentences, as illustrated in the examples in (4) below.

⁵ A *Transliteration* table is available on page 208 of this work.

- (4) a. *(Toj) Šte vidi ja.
 (he) will see her-cl.
 b. *(Toj) Šte lodkata vidi.
 (he) will boat-the vidi
 c. *Lodkata ja toj šte vidi.
 boat-the her-cl. toj will see
 d. *Lodkata toj ja šte vidi.
 boat-the he her-cl. will see

This led researchers to recognize Bulgarian as a language with a partially free word order and to look for the possible factors involved (cf. Avgustinova (1997); Dimitrova-Vulchanova & Hellan (1999); and Penchev (2005) for discussion of the various constraints on the Bulgarian word order; and Wilder & Ćavar (1994) for a comparison with Croatian and a 'minimalist' treatment of verb movement).

1.1.2 The subject – its presence and absence in the Bulgarian sentence

As already mentioned above, Bulgarian is a pro-drop or subject-null language. Therefore, the syntactic realization of the subject in Bulgarian is optional. However, the verb inflectional morphology includes subject-verb agreement suffixes, which indicate information on the subject. Inflections specifying person, number, and gender can be allocated in the main verb of the Bulgarian analytic verb forms, as well as in the auxiliaries constituting the VP, as illustrated in the examples in (5) below.

- (5) a. Šte xod**ja** na učilište.
 will go-1p.sg on school
 a' I will go to school.
 b. Šte **e** vidj**ala** vsičko.
 will be-3p.sg. seen-sg.f. everything
 b' She would have seen everything.

- c. Beše razbral istinata.
was-2p.sg. undertood-sg.m. truth-the
c' He had realized the truth.
- d. Bila zapomnila uroka.
been-3p.sg.f. memorized-sg.f. lesson-the
d' She has (supposedly) memorized the lesson.

The sentences which lack an overt subject in Bulgarian are classified into several groups (cf. Penchev, 1993). The examples presented in (4) above are all instances of the *definite personal*. The other three groups are the *indefinite personal*, the *generic personal*, and the *impersonal*, illustrated in the examples in (5a), (5b), and (5c and d), respectively.

- (6) a. Rešixa, če njama da dojdät.
decided-3p.pl. that won't to come-3.p.pl.
a' They decided that they would not come.
- b. Ti mu podaj prüst, toj šte ti otxape rükata. (Bulgarian saying)
you him give-2p.sg. finger, he will you-Dat.cl. off-bite-2p.sg. hand-the
b' Give him a finger and he will bite your hand off.
- c. Vali snjag.
rain-3p.sg. snow
c' It's snowing.
- d. Tuk se puši.
here refl.cl. smoke-3p.sg.
d' One can smoke here.

Besides *subject-null* sentences, Bulgarian displays another syntactic phenomenon sometimes called *Subject Doubling* (cf. Penchev, 1993). These are

sentences, where the subject is realized initially as a pronominal clitic followed by a lexical item or a personal pronoun, referring to the same participant, as illustrated in the examples in (7) below.

- (7) a. Te decata otidoxa na razxodka.
 they children-the went-3p.pl. on walk
 a' The children went for a walk.
- b. Tja na neja i podarixa cvetja.
 she to her her-cl. presented flowers
 b' She has been given flowers.

However, it is arguable whether the initial clitics in (7) are true reduplications of the subjects since the subject does not have to agree with the clitic, as illustrated in (7b). Alternatively, this phenomenon is referred to as the *Hanging Topic Construction* (cf. Krapova & Cinque, forthcoming). This is in line with earlier proposal in Cinque (1990) for differentiation of *Clitic Doubling* from other syntactic constructions involving overt left- or rightward movement of the clitic, as illustrated in the examples in (3) above.

1.1.3 Direct Object in Bulgarian

The direct object (O_d) in Bulgarian is realized syntactically either as a full Noun Phrase (NP) or a pronominal clitic in the Accusative, as illustrated in the examples in (6a) and (6b, c), respectively.

- (8) a. Tja otvori *prozoreca*.
 she opened window-the
 a' She opened the window.
- b. Tja *go* otvori.
 she it-cl. opened

- b' She opened it.

- c. Otvori *go*.
opened it-cl.
- c' (She) opened it.

As it was mentioned above, there are certain constraints on the syntactic position of clitics, as illustrated in the examples in (9) below.

- (9) a. *Go otvori.
it-cl. opened
- b. *Tja otvori *go*.
she opened it-cl.

In the light of the extensive literature on this matter (cf. Dimitrova-Vulchanova (1998a) and Dimitrova-Vulchanova & Hellan (1999) to mention some), I will not discuss here the possible restrictions posed on the position of the direct object clitics in the linear string in Bulgarian, as it is not directly related to this research topic.

Another issue that has been intensively researched on is commonly referred to as reduplication of the direct object or *Clitic Doubling*. As a core member of the Balkan Sprachbund, Bulgarian displays doubling of the direct object, realized simultaneously with a noun or a full pronominal form and a pronominal clitic, as illustrated in the examples in (10) below.

- (10) a. Izpix *go kafeto*.
drank-1p.sg it-cl. coffee-the
a' I drank the coffee.

- b. Vidjax *gi tjax v drugata staja*.
saw-1p.sg. them-cl. them in other-the room
b' I saw them in the other room.

Since we can not have two equivalent arguments within the same minimal clausal domain, it is assumed that the pronominal clitics do not occupy argument position when doubled. However, if realized on their own, they must be considered in argument position (cf. Penchev (1993) for discussion and analysis).

Recent research and analyses presented in Dimitrova-Vulchanova & Vulchanov (forthcoming) demonstrate that *Clitic Doubling* is attested at a very early stage of Old Bulgarian. Instances of clitic reduplication are found already in the 10th century Codex Suprasliensis (cf. Dimitrova-Vulchanova & Vulchanov (forthcoming) for discussion and references). However, at this early stage, the reduplication of the clitic is not obligatory, while in Modern Bulgarian there are cases of obligatory clitic doubling, as illustrated in the examples in (11) below with the doubling of *wh*-constituents.

- (11) a. Kogo go e strax?
 who him-cl. is fear
 a' Who is afraid?
- b. Kogoto go e strax, da ne gleda.
 whom him-cl. is fear, to not look
 b' Those who are afraid should not look.

In addition, if a specified direct object is fronted, the reduplication is also obligatory (cf. Dimitrova-Vulchanova (1998) for analysis and discussion), as illustrated in the examples in (12) below.

- (12) a. Deteto *(go) vidjax da tiča.
 child-the him-cl. saw to run
 a' I saw the child running.

It is argued that clitic reduplication in Bulgarian is related to semantic restrictions posited exclusively in the domain of Information Structure (cf. Dimitrova-Vulchanova & Vulchanov (forthcoming) on the diachrony of *Clitic Doubling* in Bulgarian and the conditions licensing it).

1.1.4 Indirect Object in Bulgarian

The indirect object in Bulgarian (O_i) is realized overtly in the sentence either as a prepositional phrase or with a pronominal clitic in the Dative, as illustrated in the examples in (13) below.

- (13) a. *Vŭrnax knigata na Angel.*
 returned-1p.sg. book-the to Angel
 a' I returned the book to Angel. / I returned Angel's book.
- b. *Vŭrnax knigata na nego.*
 returned-1p.sg. book-the to him
 b' I returned the book to him.
- c. *Vŭrnax mu knigata.*
 returned-1p.sg. him-Dat.cl. book-the
 c' I gave him back the book.

This dual expression of the indirect object in Bulgarian, either analytical (as a prepositional phrase) or syncretistic (with a pronominal clitic in the Dative) has its source in the historical process (17th – 18th century) towards analytical expression of the Old Bulgarian case system. Thus, the input from the two competing grammars has influenced the syntactic realization of arguments, visible today in the dual expression of the indirect object (cf. Mincheva (1964); Dimitrova-Vulchanova & Vulchanov (in press) for a detailed discussion on the historical evolution of the Dative in Old Bulgarian).

In addition, it must be mentioned that the sentence in (13a) has a second meaning with a possessive reading of the prepositional phrase *na Angel* (*of Angel*). This is considered a result of the process of case levelling when the Genitive possessive was substituted by the Dative. Already in Old Bulgarian the Dative clitic was used inside nominal expressions, while clitics in the Genitive were not used at all (cf. Dimitrova-

Vulchanova & Vulchanov, forthcoming). Thus, the Dative clitics in Modern Bulgarian are also used as possessive clitics.⁶

Consider the examples in (14) below.

- (14) a. Pljasnax rŭkata na deteto.
 she slapped-1p.sg. hand-the of child
 a' I slapped the hand of the child.
- b. Pljasnax rŭkata *mu*.
 slapped-1p.sg. hand-the him-Dat.cl.
 b' I slapped his hand.
- c. Pljasnax *mu* rŭkata.
 slapped-1p.sg. him-Dat.cl hand-the
 c' She slapped his hand.

However, it depends on the verb semantics whether the Dative clitic is to be considered an instance of *possessor raising* as in the example in (14c) or, it should be analyzed as clausal argument, as in the sentence in (15b) below.

- (15) a. Udarix ŝamar na momĉeto.
 hit-1p.sg. slap on boy-the
 a' I slapped the boy. (I gave the boy a slap)
- b. Udarix *mu* ŝamar.
 hit-1p.sg. him-Dat.cl. slap
 b' I slapped him. (I gave him a slap)

⁶ cf. also Dimitrova-Vulchanova & Guisti (1999) and Dimitrova-Vulchanova (2000) for analyses and comparison of English and Bulgarian possessive clitics.

In addition, Dative clitics in Bulgarian can be related to a range of semantic interpretations (cf. Petrova (2006) for detailed analysis and discussion of the semantics of Dative pronouns).

On the one hand, Dative clitics in Bulgarian are considered true arguments subcategorized for by the verb, as illustrated in the example in (16) below.

- (16) Dadox *mu* knigata. (*Recipient*)
 gave-1p.sg him-Dat.cl. book-the
 I gave him the book.

On the other hand, Dative clitics can be constructionally added (not encoded in the verb), as in the examples in (17a) and (17b) below.

- (17) a. Svarix *mu* kafe. (*Benefactive*)
 boiled-1p.sg him-Dat.cl. coffee
 a' I made him coffee.
- b. Preča *mu* da mine. (*Malefactive*)
 obstruct-1p.sg. him-Dat.cl. to pass-3p.sg.
 b' I obstruct his way.
- c. Miriše *mu* na cvetja. (*Experiencer*)
 smell him-Dat.cl. of flowers
 c' He could smell flowers.

Another instance of non-argument Dative clitic, expressing only the emotional relation of the speaker, is named *Dativus Ethicus*, exemplified in the sentence in (18) below.

- (18) Skŭp *mi* e.
 dear-m.sg. me-Dat.cl. is
 He is dear to me.

Besides, in some sentences with extended meaning of the head verb, a kind of "dummy argument" may appear on the analogy of the concrete sense of the verb. For example, compare the sentence in (15c) above with the example in (19) below.

- (19) Udarixme *mu* po njakolko čaški.
 hit-3p.pl. him-Dat.cl. each several glasses
 We drank several glasses each.

In this case, the Dative clitic is not subcategorized for by the verb. It is neither required nor provided by the clausal structure, either. In addition, it must be in third person, singular, and masculine. Therefore, it can be considered a fake argument that is analogous to the clausal argument illustrated in (15c).

Finally, like already observed with the reduplication of the direct object, a doubling of the indirect object is also displayed in Bulgarian, as illustrated in the examples in (20) below.

- (20) a. Vŕnax mu knjigata na Angel/na nego.
 returned-1p.sg. him-Dat.cl. book-the to Angel/to him
 a' I returned the book to him.
- b. Na Angel/Na nego mu vŕnax knjigata.
 to Angel/to him him-Dat.cl. returned-1p.sg. book-the
 b' I returned the book to him.

Here, we also have cases of obligatory doubling, as illustrated in the examples in (21) below.

- (21) a. Na nego *(mu) e zle.
 to him him-Dat.cl. is bad
 a' He feels sick.

- b. Na kogo *(mu) se spi?
 to whom him-Dat.cl. refl.cl. sleep-3p.sg
 b' Who feels like sleeping?

As already mentioned above, the conditions licensing the clitic reduplication in Bulgarian are argued to belong in the domain of Information Structure (cf. Dimitrova-Vulchanova & Vulchanov (forthcoming) and Dimitrova-Vulchanova & Hellan (1999) for analysis and discussion of *Clitic Doubling* and the conditions posed on it).

Having presented some of the most prominent features of Bulgarian syntax, now we can pay closer attention to verbs on the syntax-semantics interface. Since it is a contrastive study of two languages, there is the need for a brief assessment of at least few of the morpho-syntactic characteristics of English and Bulgarian verbs with an account of the information that can be lexically encoded in them.

1.2 Brief Typological Juxtaposition of English and Bulgarian Verbs

Languages display cross-linguistic variation on the interface of conceptual structure and syntax. Thus, languages may vary in event conceptualization and mapping of participants from conceptual structure to lexical items and grammar.

1.2.1 Bulgarian verbal morphology vs. English verb particles

Both English and Bulgarian are considered satellite-framed languages (cf. Talmy 1985, 1991, for a detailed discussion). Yet, the *satellites* may have different overt expressions cross-linguistically. Thus in Bulgarian, verbs are often accompanied by a prefix,⁷ while in English the *satellite* is usually realized as a particle, as illustrated by the examples in (22) below.

⁷ cf. Guentcheva (2002) for a discussion and analyses of the semantics and functions of Bulgarian prefixes.

- (22) a. Toj *otrjaza* dŭrvoto.
 he off-cut tree-the
 a' He *cut* the tree *off*.
- b. Toj *narjaza* tortata.
 he up-cut cake-the
 b' He *cut up* the cake.

On this account, though, the main verb is generally taken to encode only a supporting event, while the satellite is assumed to encode the state change.

However, based on recent research on the semantic categories of "cutting and breaking" events across languages, Majid et al. (forthcoming) argue that the meaning of the main verb inherently entails change of state, while the satellites reinforce and further specify that meaning.

Thus, accounting for the possible changes in meaning as result of the interaction of the verb satellites with the head verb, I focus on the situation that is lexicalized by the main verb and the respective participant information it encodes.

In addition, various aspectual specifications of the verbs at hand were also taken into account including completedness (telicity), duration, and definedness for end point (cf. Section 3.3.2.1 for a discussion of the terminology used in the research.)

1.2.2 Causation and Causative Alternations in English

Some of the most common lexicalization patterns in English involve the use of intransitive verbs to denote events of causation. The *Induced Action Alternation* is one of the various types of Causative Alternations described in Levin (1993). Examples of this alternation are given in (23) and (24) below.

- (23) a. The horse jumped over the fence.
 b. Angel jumped the horse over the fence.

- (24) a. The dog walked in the garden.
 b. I walked the dog in the garden.

The respective Bulgarian constructions denoting the same situations are given in (25) and (26) below.

- (25) a. Konjat skoči nad ogradata.
 horse-the jumped over fence-the
 a' The horse jumped over the fence.
- b. Angel nakara konja da skoči nad ogradata.
 Angel made horse-the to jump over fence-the
 b' Angel made the horse jump over the fence.
- (26) a. Kučeto se razhodi v gradinata.
 dog-the refl.cl. walked in garden-the
 a' The dog walked in the garden.
- b. Razxodix kučeto v gradinata.
 walked-1p.sg. dog-the in garden-the
 b' I walked the dog in the garden.

A comparison between the English examples and their respective Bulgarian counterparts shows a variation in the choice of lexical items and syntactic patterns in the lexicalization of the events presented. While the English verb *jump* can be used in the *Induced Action Alternation*, as illustrated in (23b) above, the same situation can be expressed only analytically in Bulgarian using the “make to” construction, as shown in (25b). In the case of the *walking* event, however, English and Bulgarian seem similar in lexicalizing a situation of making someone walk, as exemplified by the parallel syntactic construction in (24b) and (26b). Yet, we must notice that the Bulgarian sentence in (26a) involves the use of the reflexive clitic *se*, and therefore differs syntactically from the English example in (24a).

More examples of cross-linguistic variation in the lexicalization and the grammaticalization of similar events are given in (27) and (28), respectively.

- (27) a. The statue *stood* on the pedestal.
 a' Statujata *stoeše* na piedestala.
 statue-the stood on pedestal-the
- b. They *stood* the statue on the pedestal.
 b' Te *postavixa* statujata na piedestal.
 they set statue-the on pedestal

- (28) a. The bell rang.
 a' Zvūnecūt izzvūnja/*pozvūni.
 bell-the out-rang/*rang
- b. The visitor rang the bell.
 b' Gostūt *izzvunja/pozvūni na zvūneca.
 visitor-the out-rang/rang on bell-the

In addition, we should note that English differs typologically from Bulgarian in the prevalent employment of intransitive verbs to lexicalize transitivity events. Thus, the example in (29) below demonstrates a transitive use of *disappear*, which was thought, until recently, as impossible.

- (29) "That includes my witness, *who* you've disappeared!" (From the script of *Runaway Jury*, 2003)

In Bulgarian, however, this usage is completely unacceptable and the situation at hand can be lexicalized only by the "make to" construction, as seen earlier in the example in (25b). Thus, Bulgarian is known as having preference for the opposite process, namely, employing transitive verbs to denote intransitive events, as illustrated in the examples in the next section and the analyses of corpus data in Chapter 4.

1.2.3 *Se-Constructions in Bulgarian*

A prevailing syntactic pattern in Bulgarian involves the reflexive clitic *se* and is therefore referred to as the *se-construction*. The meaning of the Bulgarian *se-construction* varies in relation to the semantics of the verb heading the construction (cf. Dimitrova-Vulchanova (1996/99) for analysis of the possible semantic interpretations, and Koeva (2004, 2005a) for a detailed discussion on the possible transformations including *se*). Therefore, I refer to this pattern as *se-constructions* (in plural), thus reflecting the difference in their meaning and the variety of events it lexicalizes.

As suggested by the name,⁸ the *Reflexive se-construction* entails a reflexive meaning, as illustrated in the examples in (30) below, taken from the Large Corpus of Written Bulgarian (LCWB).⁹

- (30) Toj *se* plesna s dlani po čeloto ... (LCWB)
 he refl.cl. slapped with palms on forehead-the
 He slapped himself on the forehead with his palms.

Here, the reflexive clitic *se* is anaphoric, i.e. it is co-referential with the subject and marks the presence of a second participant with verbs that subcategorize for two participants.

In the *Absolutive se-construction* the subject position is occupied by the participant that is otherwise realized as the direct object of the transitive verb, as illustrated in the examples in (31).

- (31) a. Tja udari silno vratata.
 she hit hard door-the

⁸ The names used here are in line with the adopted model (cf. the discussion in section 3.3) proposed in Dimitrova-Vulchanova (1996/99).

⁹ This is the corpus I used to collect my data in Bulgarian. It is developed in the Department of Computational Linguistics at The Bulgarian Academy of Sciences, where I did my field research. I would like to thank once again the people who work there for their valuable advices and technical assistance. Detailed information on the size and the content of the corpus is given in Chapter 4 of this work.

- a' She hit the door hard.
- b. Vratata *se* udari silno. (LCWB)
door-the refl.cl. hit
- b' The door hit hard against.

The entailed meaning of the *Absolutive se-construction* is that no external agent or cause is involved in the situation at hand. This is demonstrated in the unacceptability of an eventual *by*-phrase denoting an agent or a phrase denoting a purposeful event (like "*on purpose*") and the contrasting acceptability of the manner adverbial "*on its own*," as illustrated in the sentences in (32) below.

- (32) a. * Vratata *se* udari *ot Angel*.
door-the refl.cl. hit by Angel
- b. * Vratata *se* udari *naročno*.
door-the refl.cl. hit on purpose
- c. Vratata *se* udari *ot samo sebe si*.
door-the refl.cl. hit from only oneself refl.cl.
- c' The door hit on its own.

A discussion of the semantic characteristics of the verbs which allow for the *Absolutive se-construction* is given in Chapter 4.

Another type is named the *Passive se-construction*. It differs from the *Absolutive* in that an overt realization of the participant causing the event is acceptable as illustrated in the example in (33) below.

- (33) Tortata *se* izjade *ot decata*.
cake-the refl.cl. ate by children-the
The cake was eaten by the children.

Similar to the Passive construction presented above is the *Impersonal Passive*, illustrated in the example in (34) below.

- (34) Tuk ne *se* puši.
 here not refl.cl. smoke-3p.sg.
 One cannot smoke her.

In addition, there is also a construction which is dubbed *Impersonal Optative* (Koeva 2004) and its entailed reading is the “feel like” sense (Dimitrova-Vulchanova 1996/99), exemplified in the sentence in (35) below.

- (35) Puši mi *se*.
 smoke me-Dat.cl. refl.cl.
 I feel like smoking .

As already mentioned, the differences in the meaning of the various types of se-constructions is attributed to the information encoded in the head verb in the construction at hand and the semantic characteristics of the participants in the situation denoted by the verb. The types relevant for the verbs in this research are discussed in detail along with the analyses of the corpus data presented in Chapter 4 of this work.

1.3 Conclusions

Together with an introduction of the present research topic and the sets of verbs examined, this introductory chapter discussed some of the most characteristic features of Bulgarian syntax on the syntax-semantics interface. In addition, I have included a brief juxtaposition of English and Bulgarian prevalent syntactic patterns in relation to the conceptualization of events and the mapping of conceptual categories onto lexical items and grammatical features across languages.

2. Current Relevant Approaches to the Syntax-Semantics Interface

My research intertwines with several of the existing theories in lexical semantics and the syntax-semantics interface, and while it is in line with some of their ideas, it considerably differs on others. In order to place my work among the current approaches which account for the semantics of verbs and their lexical representation in relation to the syntax-semantics interface, I must briefly describe those that I believe are closest in their agenda in relation to my work, although they may appear quite distinct or even isolated from each other.

Thus, section 2.1 presents the view of an existing interaction between verb semantics and verb syntactic behaviour as suggested in works by Levin (1993) and Levin & Rappaport Hovav (1995). It discusses the adequacy of a verb classification based on the verb meaning and its overt syntactic realization. Then some of the central ideas of Construction Grammar as presented in works by Goldberg (1995) and Fried & Östman (eds, 2004) are described in section 2.2, where I also express my own views for and against the suggestion of considering constructions as independent meaningful units on their own. Section 2.3 introduces an account for the semantic compositionality in natural languages and the generation of new word senses through the main principles of the Generative Lexicon framework as established in the seminal work of Pustejovsky (1995) and further developed in Tenny & Pustejovsky (2000) among others. Finally, in section 2.4 I bring forward the ideas of Conceptual Semantics and briefly present the organization of the tiered representation of Conceptual Structure (CS) proposed in

works by Jackendoff (1990, 1996) and Nikanne (1990, 1995). In many aspects this approach draws nearest to the lexical representation of verbs I have adopted, following the Sign Model proposal (Dimitrova-Vulchanova 1996/99) to be discussed in the next chapter.

2.1 Verb Classes and Alternations

I side with Levin in her guiding assumption (Levin 1993, p.1) that *the behaviour of a verb, particularly with respect to the expression and interpretation of its arguments, is determined largely by its meaning*. In her book *English Verb Classes and Alternations*, Levin has made an attempt to analyse the syntactic behaviour of verbs, while looking for clues for the linguistically relevant aspects of verb meaning. This preliminary investigation explores the interface of the syntactic and semantic properties of English verbs by using the set of *diathesis alternations* a particular verb allows as an indication of its semantic characteristics. Thus, she aims at categorizing the English verbs in classes according to their meaning as induced by the alternations in which they participate. The research follows a wide range of diathesis alternations, described in the first part of the book. Then verbs are grouped into classes and subclasses, and each semantic type or verb class is assumed to have a set of alternations that it participates in, and another set of alternations that are not possible with the verbs in this class.

On this approach, however, every verb may fall into several types (classes) according to the variety of alternations, in which it participates. For example, the verb *kick* is included in six rather different semantic types: *Carry Verbs*, *Throw Verbs*, *Hit Verbs*, *Split Verbs*, *Crane Verbs*, and *Verbs of Body Internal Motion*. The reason for this is that it occurs in syntactic structures common to other verbs in those classes. For comparison, the verb *strike* is included in three similarly distinct classes – *Hit Verbs*, *Amuse Verbs*, and *Verbs of Sound Emission*, and the motivation for including it in the third one is essentially the same as the reason for including *kick* into *Verbs of Internal Motion*, namely, they both display the so called *Directional Phrases with Nondirected Motion Verbs* alternation. Consider the following examples, given in Levin (1993) as instances of directional phrases with Nondirected Motion Verbs.

- (1) a. He kicked the ball *to the room*.
 b. He struck *into the back of the net*.

Both *kick* and *strike* are listed as members of the *Hit* Verbs, and while they both demonstrate the identical tendency to appear with directional phrases, *kick* is assumed to belong to Verbs of Body Internal Motion and *strike* is considered a Verb of Sound Emission on the same grounds.

What about the verb *hit* then? According to Levin, it is a member of the classes of *Hit* Verbs, *Throw* Verbs, Non-Agentive Verbs of Contact by Impact, and Verbs of Contiguous Location. Even though on Levin's classification *hit* is not considered a directed motion verb, still the following example, taken from the British National Corpus (BNC), shows that *hit* can appear with a directional phrase.

- (2) ... it is hit *out of the way* ... (BNC: GVF 1898)

Is *hit* then to be considered a verb of Body Internal Motion or a verb of Sound Emission? Or even belonging to yet another class?

At this point Levin's Verb Classification does not meet the expectations originally presupposed by the approach at hand. Instead of planting multiple cross-references of verb classes and alternations, we need to seek and distinguish the subtle semantic features of the verbs that are in the basis of the patterns of syntactic behaviour they display.

I find equally problematic classifying *hit*¹ and *strike* as members of the *Amuse* Verbs class, a subclass of the more general class of *Psych* Verbs, also referred to as Verbs of Psychological State. We do come across occurrences of these verbs, which appear to refer to psychological states as exemplified in the sentences taken from the BNC and presented in (3) below.

¹ Although not enlisted as a class member, *hit* is mentioned in the comments to chapter 31.1 *Amuse* Verbs (Levin, 1993, p. 191) as one of the verbs in this class, together with *strike*, that *also are used in physical action sense*.

- (3) a. *The unfriendliness in her voice* struck Emmie like a jet from a cold hose. (BNC: HH9 1535)
- b. As I was going up to my cell *it all* hit me; am I ever going to get my daughter back? (BNC: FR5 2433)
- c. ... he was suddenly struck *by inspiration*. (BNC: A7H 681)

Rather than looking for a class that is suitable for this usage of the verbs, I suggest that we consider the fact that verbs can head sentences with extended non-physical action sense as a result of an ability to take referents, whose denotational properties are perceived as being in conflict with the type of situation originally denoted by the verb. This ability is most likely a result of the general properties of syntax. That is, we are looking into cases of compositional semantics (discussed more explicitly in sections 2.2, 2.3 and 4.3), whereby non-canonical participants trigger the preferred sense or reading of the verb at hand. In addition, this may have an effect on the verb subcategorization preferences (cf. Hare et. al (2003) for a similar observation addressed in section 3.2).

While Levin's research does help *pave the way toward the development of a theory of lexical knowledge* (Levin, 1993, p.1), it does not follow entirely her initial agenda. Both Levin (1993) and Levin & Rappaport Hovav (1995) use alternation patterns as a valuable tool to map out ways in which verbs naturally group together. However, this methodology is not reflected in the actual verb classification. At least in this preliminary investigation, verbs are grouped and classes are named in a somewhat introspective way that fails to avoid cases of including one verb into several rather different classes. Although in subsequent work (Levin & Rappaport Hovav, 1995) close attention is paid to verb semantics with respect to its effect on the syntactic realization of the verbs at hand, it is still not present in the actual verb grouping, which is left exactly the same. Yet again, it is doubtful whether any word enumerative classification would capture the subtle nuances of verb semantics. Instead, I would suggest that a more net-like distribution should be considered, whereby verbs are grouped according to the types of situations they can denote as visible from their syntactic behaviour (more detailed discussion follows in Chapter 4, section 4.2 of the present work).

2.2 Construction Grammar

One of the frameworks that explore the nature of verb meaning in its relation with sentential meaning is *Construction Grammar* (Fillmore 1988, Kay 1990, Lakoff 1987, Goldberg 1995, Fried and Östman (eds) 2004, among others).

The Construction grammar approach has grown largely out of works on *Frame Semantics* (Fillmore 1975, 1985) and its crucial belief is that constructions, as form-meaning correspondences, carry meaning independently of the words in the sentence and that some specific semantic structures together with their associated formal expression must be recognised as constructions independent of the lexical items which instantiate them. Furthermore, constructions are assumed to exist independently of particular verbs. Yet, it is admitted that the interaction between verb meaning and constructional meaning is a complex one, where top-down and bottom-up processing occur simultaneously. That is, the semantics of the verb adds to the semantics of the construction and vice versa - the construction does not simply impose its meaning on the verb. Support for such an interactive mechanism has been found in its successful implementation in connectionist models already in 1986 (Rumelhart & McClelland, 1986).

Within the theory of Construction Grammar, it is also assumed that the basic means of clausal expression in a language is provided by a special subclass of constructions called *argument structure constructions*. Examples of English argument structure constructions (Goldberg 1995) include the following:

Ditransitive	X CAUSES Y TO RECEIVE Z	Subj V Obj Obj2	Pat faxed Bill the letter.
Caused Motion	X CAUSES Y TO MOVE Z	Subj V Obj Obl	Pat sneezed the napkin off the table.
Conative	X DIRECT ACTION at Y	Subj V Oblat	Sam kicked at Bill.

One of the strongest arguments for the existence of these constructions is seen in the way they can account for the appearance of a verb in many different alternation

patterns. An example (in Goldberg, 1995) is given with the verb *kick* participating in eight distinct argument structures:

1. Pat kicked the wall.
2. Pat kicked Bob black and blue.
3. Pat kicked the football into the stadium.
4. Pat kicked at the football.
5. Pat kicked his foot against the chair.
6. Pat kicked Bob the football.
7. The horse kicks.
8. Pat kicked his way out of the operating room.

The constructionalist approach to this problem is opposed to all other linguistic theories² that *attempt to predict overt syntax on the basis of semantic roles or theta role arrays*. According to Goldberg, they treat the verb as an *n*-place relation that expects the exact number of arguments of the correct type to fill in the empty slots and they must posit a new sense on the verb to explain the existence of every single syntactic configuration. In Construction Grammar, this problem is claimed to be handled by accounting for the interaction between verb meaning and construction meaning.

However, in order to explain this interaction, Construction Grammar also posits roles as semantically constrained relational slots in the dynamic scene associated with the construction. A distinction is made between *participant roles* (delimited by the verb's semantics) and *argument roles* (associated with the construction). This division is seen as nearly parallel to the difference in Dowty's (1986) earlier treatment of *individual thematic roles* versus *thematic role types*. This differentiation intends to capture the fact that verbs are associated with frame-specific roles, whereas constructions are associated with more general roles such as *agent*, *patient*, and *goal*. Thus, participant roles are viewed as instances of the more general argument roles and are expected to reflect specific selectional restrictions as well, which is one of the basic

² Where only the Government and Binding Theory (Chomsky, 1981) and Lexical-Functional Grammar (Bresnan 1982) are explicitly mentioned by Goldberg (1995).

principles in Fillmore's Frame Semantics. Therefore, when a verb is associated with a construction, the participant roles of the verb may be semantically *fused* with the argument roles of the construction, where fusion is meant to signify the simultaneous semantic constraints on participant roles and argument roles. The possibility of roles fusing is therefore determined by the compatibility of their types and the rules of fusion are determined by two principles, presented below as given in Goldberg (1995).

1. *The Semantic Coherence Principle*: Only roles, which are semantically compatible, can be fused.
2. *The Correspondence Principle*: Each participant role that is lexically profiled and expressed must be fused with a profiled argument role.

Thus, the representation of an argument construction is assumed to consist of a pairing between a semantic level and a syntactic level as shown in Fig.1 below.

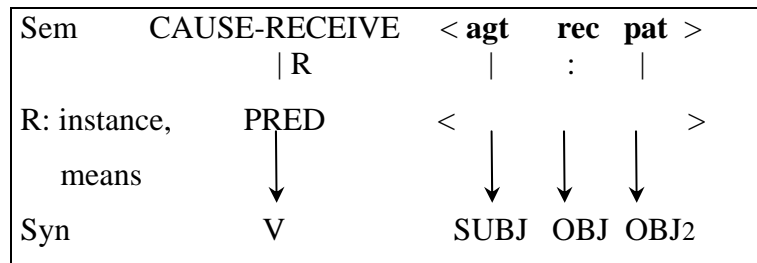


Fig.1 Ditransitive Construction

The semantics associated directly with the Ditransitive Construction is 'CAUSE-RECEIVE <**agt pat rec**>.' PRED is a variable that is filled when a particular verb is integrated into the construction. The roles indicated by solid lines are obligatorily fused with participant roles, whereas the roles indicated by a dashed (dotted) line are not obligatorily fused with roles of the verb, that is, they can be contributed by the construction. The type of relation R specifies the way, in which the verb is integrated into the construction.

Having said this, it turns out that Construction Grammar also relies heavily on semantic roles to judge on the possibility of a verb to enter a particular semantic

construction, and thus to appear in one syntactic pattern or another. The main difference then between Construction Grammar and the rest of the linguistic theories is that on the constructional approach, the different argument structures are viewed as constructions and assumed to have independent meaning.

Instead of positing semantic roles on the participants for each syntactic structure in which a particular verb may appear, I suggest that we look at the situation that a verb can lexicalize as a whole, paying attention to its sub-events and how the semantics of the verb is thus expressed overtly. This view is closer to the fundamentals of the constructionist approach - Frame Semantics, since according to Fillmore (1977) *meanings are relativized to scenes*.

Therefore, it is hard to believe that constructions alone carry the intended meaning, i.e. that they exist as disconnected meaningful units. However, I consent to the constructionist approach to the extent that I consider some constructions may exist on the border between lexicon and conceptual structure, as coined expressions, with semi-idiomatic meaning, which is usually marked by a very specific phrase (as in the *Way*-construction) or a certain type of clause (as in the *Resultative* construction). This is also in line with Jackendoff (1990) who regards the *Way*-construction as a '*constructional idiom*.'

Thus, I explore the existence of constructions mostly in the way they are presented by Nikanne in Chapter 7 of Fried and Östman's *Construction Grammars* (eds, 2004), where constructions are introduced as a separate module (apart from the lexicon) which consists of combinations of syntactic and conceptual structures, and includes some pragmatic information. In addition constructions may contain a specific lexical item as the word *way* in the *Way*-construction to be discussed in section 2.2.2 below. In order to account for this view, I will briefly discuss two of the constructions³ that I have mentioned so far, namely the *Resultative* construction and the *Way*-construction.

³ To avoid terminological confusion, I will continue to use the term construction when referring to the structural units proposed by Goldberg (1995).

2.2.1 More on the Resultative construction

The first thing to be said about the resultative construction is that it is highly constrained semantically. Only participants that may undergo *a change of state* (Goldberg 1995) as a result of the action denoted by the verb can appear with resultatives. Traditionally, these participants are defined as *patients* (Lakoff 1976), which is also the assumption in Construction Grammar. On the terminology used in my work (Chapter 3, section 3.3), I will refer to these participants as being marked with the value of *Monodeveloper* (cf. also Dimitrova-Vulchanova 1996/99 for a detailed analysis). This distinction is important in the analysis adopted here, as it will account for a unified treatment of resultatives.

Resultatives are overtly marked by a phrase (usually an adjective phrase as in the example in (4a) and (4b) or a prepositional phrase as in the example in (4c)), which spells out the resulting state of the participant overtly realized by the predicative.

- (4) a. She ... pulled her sash *free* ... (BNC: H8J 1860)
 b. ... We were all struck *dumb* for the moment. (BNC: B3F 575)
 c. ... Billy stabbed Miller *to death* ... (BNC: GT4 182)

As the resultative construction is also found with intransitive verbs, Goldberg (1995) argues that the so-called fake object (named by Simpson 1983) need not be a real argument of the verb, but is instead provided by the construction as seen in Fig.2 below.

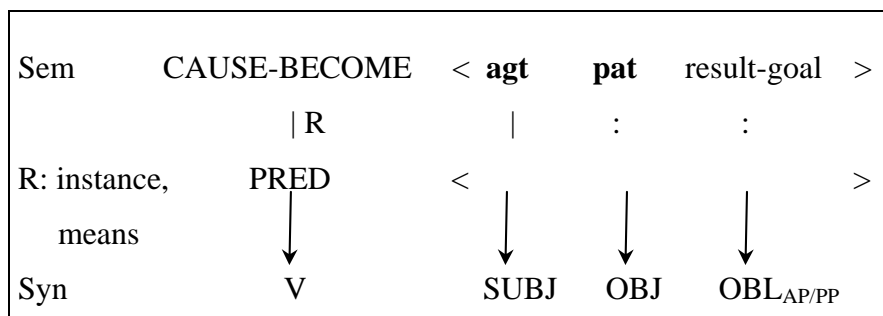


Fig.2 Resultative Construction

Although it seems to account for the grammaticality in sentences like the one in (5a) and to rule out ungrammatical sentences like that in (5b) and (5c), this approach still does not answer the question why sentences like the ones in (5d) and (5e) are ungrammatical.

- (5) a. He talked himself hoarse.
 b. *He talked hoarse.
 c. *He talked himself.
 d. *He talked her hoarse.
 e. *She laughed him silly.

In addition, the construction approach does not account for the cases like the example in (6a) and (6b) in a way consistent with the rest of the theory.

- (6) a. The pond froze solid.
 b. The toy broke apart.

Thus, a different, two-argument construction is posited to exist with respect to intransitive resultatives as shown in Fig. 3 below.

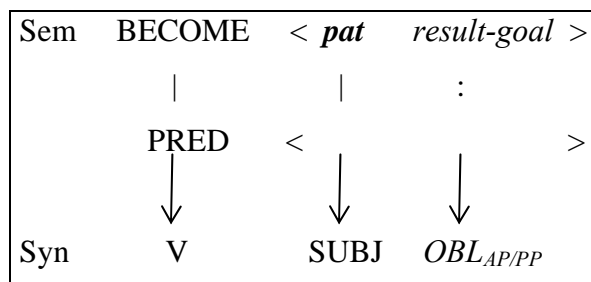


Fig.3 Intransitive Resultative Construction

Even though Goldberg does not see it as a drawback of the proposal, this indicates a theoretical contradiction. Thus, one "meaning unit" (which should be equal to one construction) is expressed in two different constructions. Moreover, defining the

argument as 'patient' misses an important generalization about the properties of this participant.

Therefore, I will argue that not all resultatives are real constructions in the Construction Grammar sense. This is in line with a recent proposal by Dimitrova-Vulchanova (2003) where a unified analysis of resultatives is suggested, based on the differentiation between connected (conservative) results and disconnected (radical) results and the respective lexical constraints that govern the appearance of a verb in the two distinct types of resultative construction.

Following the semantic constraint stated in the beginning of this section, I will assume that to allow for a resultative predicate the sentence must contain an argument with the value *Monodevoleper*, which is the case with transitive verbs denoting an event that affects the participant overtly expressed as direct object (or as the subject of a passive). This view accounts for the grammaticality of some transitive verbs while naturally excludes others like *watch* and *believe* used in Goldberg's (1995) examples and reproduced respectively in the examples in (7a) and (7b) below.

- (7) a. *He watched the TV broken.
 b. *He believed the idea powerful.

Neither *watch* nor *believe* encodes a participant with the value *Monodeveloper*, in their lexical representation. Therefore, according to the constraint just introduced, it was never expected that these verbs can enter the Resultative construction, which is reflected in the ungrammaticality of the examples in (7). In fact the sentences in (7) are not only ungrammatical, they are conceptually incomprehensible, as the situations denoted by these verbs do not imply any subsequent changes that the participant realized as the direct object in the sentence may undergo, thus receiving the value of *Monodeveloper* and becoming a possible candidate for participation in a Resultative Construction.

In cases when a resultative is applied to an intransitive verb, there are two options. If the participant expressed as the syntactic subject carries the value of *Monodeveloper*, it directly meets the requirement to take a resultative phrase as in the examples in (6a) and (6b) above, and the example in (8).

- (8) *The window* smashed to pieces.

On the classification made by Dimitrova-Vulchanova (2003), all these sentences display connected results, i.e. this is a natural outcome of the situation denoted by the verb. This conforms with the analysis proposed here, namely that this type of resultative constructions is the syntactic expression of a specific semantic feature, encoded in the lexical representation of the verbs at hand.

In cases of disconnected resultatives, however, we have two distinct and not immediately connected events which we want to engage in a logical cause-result relation (cf. Pustejovsky's event structure (Pustejovsky, 1995), discussed also in section 2.3.1.2 of the present chapter). In syntax these two events are expressed by the main verb (the cause) and the small clause (the result).⁴ Thus, for the sentence in (5) above, we deal with two separate events – a) He talked, and b) His voice is hoarse, and we want to relate the state of *him being hoarse* to the situation expressed by the head verb, namely the *talking*. Thus, the co-indexation is not only syntactically but also semantically the only option. Therefore, I will suggest that the reason for the appearance of a fake object is as much semantically as it is syntactically justified. Since the subject of the verb *talk* is not inherently specified as a *Monodeveloper*, a fake object is introduced in the canonical position of the affected participant, thus satisfying the constraint. In case of (5a) this must be a reflexive pronoun, since there is no logical way of connecting an eventual result of *her being hoarse*, as the example in (5d) with the event of *his talking*.

However in the examples in (9a) to (9c) below, co-indexing of a fake object with the subject of the sentence is not necessary. This can be based in the fact that although the result is not connected to the situation denoted by the verb within the verbs lexical representation, a logical connection between the cause event and the resulting event can

⁴ A very good syntactic explanation is given in Dimitrova-Vulchanova (2003) which accounts for the syntactic realization of resultatives as a result of the unification of the two structures attained through a co-indexation of the subject of the resultative small clause with the O_d position in the head verb structure, where the cases of verbs that do not subcategorize for a complement are handled by means of embedded structures.

be easily established on the basis of pragmatic information inheritance. Consider the examples in (9a) through (9c).

- (9) a. The dogs barked us awake.
 b. The sound of footsteps on the deck above her head brought her fully awake. (BNC: H7W 2548)
 c. 'Dad,' said Sam, drumming him awake with blows on the face ... (BNC: FSP 1633)

Although neither of the verbs *bark*, *bring*, nor *drum* does include such a result in its semantic representation, language users can easily establish a logical connection between the event denoted by the verb at hand and the respective state denoted by the adjective as occurring as a result of that event. Thus, I consider the Resultative construction as a means of natural languages to establish a cause-result relation between two usually unrelated events which is also supported by the syntactic explanation given in Dimitrova-Vulchanova (2003). However, whether a language has this option available in its grammar is subject to cross-linguistic variation.

2.2.2 *The Way-construction – Constructional Polysemy?*

While Goldberg regards the resultative construction as essentially one semantic unit with two syntactic expressions, the *Way*-construction has the opposite problem – it is treated as one structural unit with two different meanings.

The formal syntactic representation of the construction as given in Goldberg (1995) can be seen in (10) below, where V is held to be a non-stative verb, and OBL stands for a directional phrase (in my work this is referred to as the *Path* component).

- (10) [SUBJ_i [V [POSS_i way] OBL]]

The *POSS_i way* in (10) above represents a phrase consisting of a possessive pronoun defining the lexical item *way* which is a constant, that is, its presence is obligatory for

this construction and it does not change. As the sameness in indexing suggests, the possessive pronoun agrees in person, number, and gender (where applicable) with the subject of the sentence. Thus, the overt syntactic realization of the *Way*-construction is quite constrained. However, two distinct meanings are ascribed to it and these are discussed in detail in the next two sections.

2.2.2.1 The Means Interpretation

Following Jespersen (1949) on the analysis of the *POSS_i way*-phrase as a type of "object of result," Goldberg (1995) argues that the means interpretation of the *way*-construction implies the *creation* of a path, either physical or metaphorical, as a result of the action denoted by the head verb. The action is also believed to be carried out in spite of some external difficulty. The examples in (11) from the British National Corpus may be considered as having Means interpretation.

- (11) a. The raiders smashed their way *into the trailer* ... (BNC: CBF 3030)
 b. Turning, he pushed his way *to the door*. (BNC: CR6 498)
 c. Himself, he would have died fighting for his life, clawing and scratching his way *out of existence*. (BNC: GUG 1398)

The construction is then represented in general as shown in the example in Fig. 4 below, where the **createe-way** role and the **path** role are provided by the construction, thus leaving for the verb to enter the construction with *only one obligatory argument*,⁵ namely the subject which must be construed as a type of **creator-theme**.

⁵ This is not consistent with the lexical representation of any of the verbs in the examples in (11) above, since they are all transitive and encode a participant, which although not overtly expressed, absorbs the force released by the action denoted by the main verb (cf. Chapter 4 for detailed analysis of the information that is lexically encoded in this type of verbs).

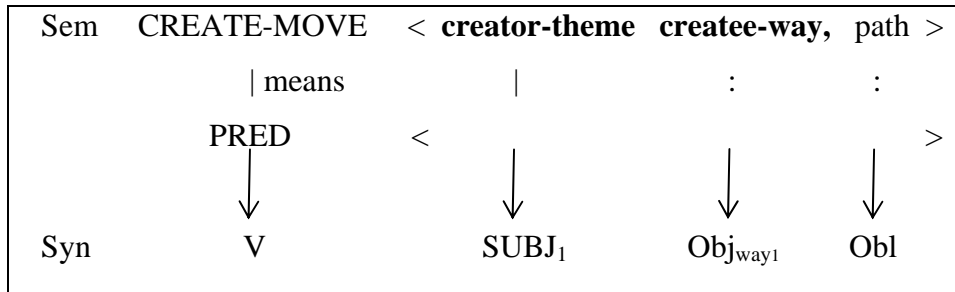


Fig. 4 *Way* Construction with Means Interpretation

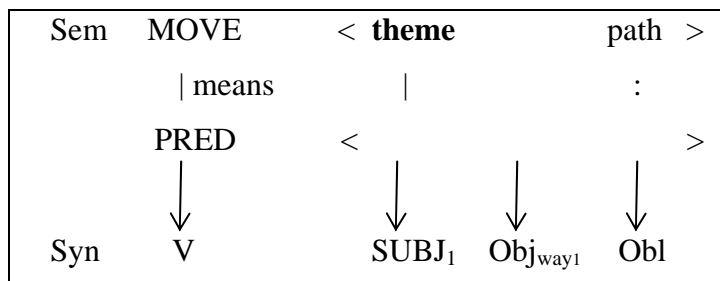
The *means* link between the CREATE-MOVE predicate and the verb (indicated by PRED) is said to represent the possibility of the verb to encode *the means of effecting motion through self-created path* (Goldberg 1995).

2.2.2.2 The Manner Interpretation

Cases, which do not seem to indicate means interpretation nor involve motion through a difficult path, are interpreted as attesting Manner instead. The examples in (12) below were also found in the British National Corpus.

- (12) a. ... Anthea tapped her way *out of the room*, ... (BNC: C8D 2772)
 b. A barge slapped its way *along the shimmering river*. (BNC: G1W 1538)
 c. A decrepit old man, supported by a young boy in a huge coolie hat, was tapping his way *down the alley*, patterned robes trailing in the rainwater. (BNC: GVL 17)

In this case the construction is thought to represent a two-place relation where the *way*-phrase is not part of the semantics of the construction itself, but is "*instead encoded as syntactic stipulation about the form of the direct object complement*" (Goldberg 1995, p.210), as it can be seen in Fig. 5 below.

Fig. 5 *Way* Construction: Manner Interpretation

This is at least challenging for the overall agenda of the construction theory and in particular recognizing one syntactic construction and distinguishing two separate meanings results in the following contradiction: while in the Means interpretation the *way*-phrase is provided by the construction, in the Manner interpretation it is just the opposite, thus arguing that the verb enters the *Way*-construction and provides the *way*-component.⁶ Straightforwardly, these are either two different constructions (which they are obviously not syntactically) or one construction with one basic meaning, which can be extended depending on the semantics of the verb entering the construction. I will argue for the latter case, using as evidence examples provided by Goldberg (1995).

2.2.2.3 The etymology of the way construction – the diachronic evidence

In order to decide which of the two senses is more basic, Goldberg (1995) follows the etymology of the *Way*-construction. The diachronic investigation is presented through the examples cited in (13) below.

- (13) a. *I **made** my way ... unto Rome.* (OED, 1400)
 b. *[He] **hew'd** out his way by the power of the Sword.* (OED, 1694)
 c. *The muffin-boy **rings** his way down the little street.* (OED 1836)

⁶ If we can freely stipulate about the form of the direct object complement, what would be the reason for positing a construction in the first place.

The first appearance of *POSS way*-phrase with a path has been first attested with the verb *make*, which suggests that in fact the *creation-of-path* reading antecedes the manner reading and is therefore more basic. In my opinion this is exactly the case.

Diachronically it appears that the phrase *make one's way* *Obl_{path}* (which is considered now as a separate idiom) has been gradually used with larger and larger sets of verbs. Initially, only verbs affecting the participant on which the action is performed could enter the construction. This is due to the implication they hold that they have an effect on the participant⁷ expressed by the direct object, thus allowing for a reading very similar to the one expressed by *make* in *make one's way*. In addition they enrich the semantics of the event (making of a path) with their own semantics. Thus, the action denoted by the verb at hand is realized as the means by which this path was created.

As the basic sense of the phrase *make one's way* has already been used in extended sense, it becomes easier to apply verbs that will exploit the central meaning of the phrase in an increasingly wider sense. The first such example has been attested in 1836 with the verb *ring*, which is twice as fast as it took for a verb different than *make* to be used in the first place. This suggests that *make one's way* has turned into productive phraseological unit with a semi-idiomatic meaning. That is, the construction is the same, but the meaning to a great extent is provided by the verb at hand. Thus, the semantics of the verb entering the construction skews the interpretation to providing the means in creating a path or the manner by which the movement along a path has obtained.

The cases of semantic ambiguity within the sentence (examples from Jackendoff 1990) may also be considered as pointing towards the idea that the two senses are not completely distinct but instead should be considered in a semantic continuum.

- (14) a. Sam joked his way into the meeting.
 b. He belched his way out of the restaurant.

The analysis proposed by Jackendoff for the sentences in (14) demonstrates the blend of the two interpretations as he suggests that they can be interpreted both with

⁷ This is the participant with the value of *Absorber* in the terminology used in this work (cf. section 3.3).

Means and Manner readings. This is also in line with current research on Manner encoding (Martinez & Dimitrova-Vulchanova, in press), where besides the critique on considering the notion of Manner as a primitive, it is also suggested that Manner should be regarded as a bundle of features which also includes Means among other.

This, however, leaves us with the question of the semantic constraints on the verbs which may appear in the construction. According to Goldberg (1995) the verbs must denote an unbounded activity (following also Jackendoff 1990), which involves self-propelled motion in a certain direction. If this was enough, verbs like *go*, *move*, and *run* should be the perfect candidates, while *walk*, for example, should not be acceptable. According to the examples in (15) and (16),⁸ however, this is not the case.

- (15) a. *I went my way to Australia.
b. *I ran my way to the city.

- (16) The old man walked his way across the country to earn money for charity.

The grammaticality of the sentence in (16) is explained in Goldberg (1995) with the implication of "motion despite difficulty." This, however, does not explain the fact that the same sentence is unacceptable with *go* as we can see in the example in (17).

- (17) *The old man went his way across the country to earn money for charity.

On my account, the acceptability of *walk* versus the unacceptability of *go* additionally proves the hypothesis that the verb entering the so-called "way-construction" must encode a manner feature, at least with comparison to a more basic/general verb. Thus the English *walk* encodes for [+on foot] when contrasted to *go*.

This can be also seen in the other example provided by Goldberg (1995) as acceptable with the verb *walk*, cited in (18) below.

- (18) The novice skier walked her way down the ski slope.

⁸ This particular sentence is cited in Goldberg (1995).

Although Goldberg discards the possibility of manner interpretation because "these verbs do not code any salient manner" (ibid. p.205), I find it even harder to believe that the sentence in (18) can be interpreted as motion despite difficulty. It is obvious that for a novice skier it would be easier to walk down the hill instead of ski down the hill. Therefore, I argue that *walk* is acceptable precisely because it encodes the [+on foot] feature, as opposed to skiing in this particular case.

Finally, even if some constructions seem to exist as separate meaningful units, we should not straightforwardly conclude that they are completely independent of the lexical items which instantiate them.⁹ In addition, we should also consider the fact that verbs denote situations and situations of one and the same type must employ similar participants, expressed similarly in the language. Thus, verbs which originally lexicalize one type of event may be forced to lexicalize another event type as seen with the *way*-construction, while preserving and extending their basic meaning.

2.3 The Generative Lexicon

Although quite different on the surface, my research shares some of the main ideas of *The Generative Lexicon* (Pustejovsky, 1995) – a framework for the semantic analysis of natural language. It does not treat the lexicon as a list of stationary word senses, entered as separate lexical items and tagged with syntactic, morphological, and semantic information. Much in the tradition of generativity in syntax,¹⁰ *The Generative Lexicon* accounts for the generativity of word senses in natural language. That is, it aims at postulating lexical primitives, which should be able to derive a potentially infinite number of senses for words from finite resources, thus explaining the interpretation of words in context and accounting for systematic relatedness between word senses in a formal and predictable way.

⁹ Similar to Goldberg's view of the constructions is Borer's radical syntax approach (cf. Borer, 2004) where she argues based on data from children language acquisition in Hebrew that the argument structure is independent from vocabulary, i.e. that lexical items do not affect the syntax of argument structure.

¹⁰ Cf. seminal works by Chomsky (1957, 1965) among others.

2.3.1 The lexicon as a multilevel system of semantic representations

The generative lexicon as seen in Pustejovsky (1995) is a computational system which consists of at least four different levels of semantic representation.¹¹ The number and type of arguments that a lexical item is considered to carry are specified in the *argument structure*. The event type of a lexical item or a phrase is defined in the *event structure*, which in addition characterizes the internal, sub-eventual structure of the lexical item. The different modes of predication, possible with a lexical item, are represented in a *qualia structure*; and the relation of a particular lexical structure to other structures in the dictionary is identified in the *lexical inheritance structure*. Thus, the semantics of a lexical item is defined as a structure representation that consists of four components. These are dubbed *Argument structure*, *Event structure*, *Qualia structure*, and *Lexical Inheritance structure* and are briefly presented in the following four sections 2.3.1.1 to 2.3.1.4, respectively.

2.3.1.1 Argument structure

Argument structure is seen as syntax independent minimal specification of the lexical semantics of a word. Within this structure, four distinct types of arguments are distinguished: *true arguments*, *default arguments*, *shadow arguments*, and *true adjuncts*.

The elements that are obligatorily expressed in syntax are defined as *true arguments*, while those arguments that are logically essential, but may be left overtly unexpressed in syntax are called *default arguments*. The difference between them is demonstrated in the different status of *John* and *the window* in the example in (19a) and respectively *Mary* and *the doll* in (19b), as opposed to *wood* from (19b) below.

- (19) a. John broke *(the window).
 b. Mary carved the doll (out of *wood*).

¹¹ This is very much in line with tiered representation (cf. Jackendoff 1990, Nikanne 1990) to be presented in section 2.4 below.

Shadow arguments refer to those arguments that are part of the semantic content of the lexical item, but their overt syntactic expression is possible only under specific conditions within the sentence. Under these conditions, the actual expressed argument must be a subtype of the shadow argument itself as shown in the example in (20).

(20) John kicked the fence with his right leg/*with his leg.

Shadow arguments could also be seen as a subclass of default arguments, but due to the specificity of the constraints on their overt syntactic expression, they are distinguished as a separate logical type.

The parameters which are not connected to the semantics of any particular lexical item but are part of the situational interpretation are called *true adjuncts*. Temporal expressions like *last week* and locative modifiers like *in Rotherham* in the example in (21) are typical cases of true adjuncts.

(21) ... James Doy was shot dead last week in Rotherham ... (BNC: CH6 735)

Along with the refined distinction between argument and adjunct phrases, the importance is placed on the compositional operations which may affect the logical status of a phrase as a certain argument type. This is demonstrated in the examples in (22) (used in Pustejovsky 1995), where the true argument *to John* in (22a) is *defaulted*¹² by the semantics of the complement in (22b) and hence becomes an optional argument.

(22) a. Mary showed her painting to John.
 b. Mary showed a movie (to John).

Thus, the argument structure of a lexical item has the following abstract representation as shown in Fig. 5 below.

¹² Pustejovsky's term used to denote a true argument which has been turned into a default argument.

$$\left[\begin{array}{l} \alpha \\ \text{ARGSTR} = \left[\begin{array}{l} \text{ARG}_1 = \dots \\ \text{ARG}_2 = \dots \\ \text{D - ARG}_1 = \dots \\ \text{S - ARG}_1 = \dots \end{array} \right] \\ \dots \end{array} \right]$$

Fig. 5

The type of the argument is directly encoded in the argument structure with $\text{ARG}_1, \dots, \text{ARG}_n$ representing true arguments, D-ARG being a default argument, and S-ARG standing for a shadow argument. The logical distinction in argument types is thus expected to play an important role in the mapping from semantic form to overt syntactic expressions.

2.3.1.2 Event structure

The Event structure level represents the event associated with the lexical item not as a single variable but as a relation between the event and its proper sub-events. Thus, it accounts for the underlying semantics of the sub-events constituting the main event and how it is projected in syntax.

Events are sub-classified into at least three sorts: PROCESSES, STATES, and TRANSITIONS. The internal configuration properties of events are also taken into consideration. It is assumed in Pustejovsky (1995, p. 73) that *events have at most a binary event structure and that there are three temporal ordering relations realized in language*. These temporal ordering relations are stated as "exhaustive ordered part of," $<_{\infty}$ (when the two sub-events are temporally ordered and the first one precedes the second, like in *build*), "exhaustive overlap part of," \circ_{∞} (express a relation of two completely simultaneous sub-events, as in *accompany*), and "exhaustive ordered overlap," $<^{\circ}_{\infty}$ (the relation of two otherwise simultaneous sub-events, but the first one starts before the other, as in *walk*). In addition, the notion of *event headedness* is introduced to reflect the part of the matrix event that has received the focus by the particular lexical item as for example illustrated in the difference between *buy* and *sell*.

Thus, event structure is thought to have an abstract representation as shown in Fig. 6 below.

$$\left[\begin{array}{l} \alpha \\ \text{EVENTSTR} = \left[\begin{array}{l} E_1 = \dots \\ E_2 = \dots \\ \text{RESTR} = \dots \\ \text{HEAD} = E_i \end{array} \right] \\ \dots \end{array} \right]$$

Fig. 6

Taking these three parameters into account, namely the binary event structure (given as E_1 and E_2 in the Fig.6), the three possible temporal ordering relations (entered under RESTR), and the headedness (including the possibility of unheaded and double-headed constructions), Pustejovsky determines the twelve event structures possible that may give rise to the different verb types. This is very close to my standpoint, since I also treat verbs as lexicalizing different event types, which is naturally reflected in their lexical semantic representation.

2.3.1.3 Qualia structure

Qualia structure is taken to specify the word's meaning in four different aspects: CONSTITUTIVE, FORMAL, TELIC, and AGENTIVE, which together shape our perception of an object or a relation in the world. Although every category is expressed with a qualia structure, not all four roles need to be present in a particular lexical item.

The CONSTITUTIVE quale reflects the relation between an object and its proper parts. This includes not only physical relations expressed in *material* or *weight*, but also any logical part-of association, which relates to *parts and component elements* (as for example the CONS quale of *hand* includes the relation [part_of(x, y:body)]).

The FORMAL quale employs the essential characteristics of an object which distinguish it within a larger domain. These include *orientation*, *magnitude*, *shape*, *dimensionality*, *colour*, and *position*. If the information contributed by the FORMAL quale is sufficient to restrict the type of the argument, it is regarded as *simple typing*,

and the value of FORMAL role is taken as identical to sortal typing of the argument. For items of *complex (or dotted) type* the FORMAL quale provides information on how arguments are related to each other, but additional constraints must be posited to avoid the generation of unattested complex types in a language.

The TELIC quale defines the purpose or the function of the lexical item. By incorporating two distinct modes, namely *direct telic* and *purpose telic*, the TELIC quale captures aspects of several different θ -roles, without being associated with any of them exclusively. Thus, the problematic one-to-one mapping between θ -roles and qualia is avoided.

Direct telic relates to items which are directly acted upon as the noun *beer* in the example in (23) adopted from Pustejovsky (1995).

$$(23) \left[\begin{array}{l} \text{beer} \\ \text{ARGSTR} = [\text{ARG}_1 = x : \text{liquid}] \\ \text{QUALIA} = \left[\begin{array}{l} \text{FORMAL} = x \\ \text{TELIC} = \text{drink}(e, y, x) \end{array} \right] \end{array} \right]$$

Purpose telic describes objects that are used in performing an activity, as for example instruments as illustrated by the noun *knife* in (24) below.

$$(24) \left[\begin{array}{l} \text{knife} \\ \text{ARGSTR} = [\text{ARG}_1 = x : \text{tool}] \\ \text{QUALIA} = \left[\begin{array}{l} \text{FORMAL} = x \\ \text{TELIC} = \text{cut}(e, x, y) \end{array} \right] \end{array} \right]$$

Thus, *telic quale* is seen as directly encoding semantic relations which are mapped onto possible syntactic alternations. For example, it accounts for the *Instrument Subject Alternation* (Levin 1995) by positing that agents and instruments share the same causative structure.

The AGENTIVE quale spells out the factors involved in the "bringing about" of an object. These include *creator*, *artefact*, *natural kind*, and *causal chain*. The AGENTIVE quale plays role in the licensing of the "sense in context" phenomenon (Pustejovsky

1995), which allows a lexical item to have one lexically specified sense that may result in different overall effects. A typical example is the verb *bake*, which when used with objects of natural kind (like *potatoes* for example) has simply the sense of changing the internal state of the object, while if used with an artefact (like *a cake*) acquires the sense of a creative activity. This is due to the semantics of artefacts which specifies their origin as a result of an activity. Thus, the AGENTIVE quale is seen as accounting for the semantic generalizations that natural languages may express.¹³

2.3.1.4 The lexical inheritance structure

The Lexical Inheritance structure contributes to the integration of the other three levels, described above. Thus, the argument, the event, and the qualia structures are brought together to jointly form lexical semantic representations.

The interaction of the semantic levels is achieved through a system based on typed feature structures (cf. Carpenter 1992, 1997 for theoretical settings and Copestake et al. 1993 for application to lexical semantics). This system is seen as consisting of two parts – a type hierarchy and a system of constraints to which types are confined. The four levels of representations are incorporated into this typing system, thus accounting for generativity and the creative use of natural language.

2.3.2 *The generative semantic operations*

The compositional interpretation of words in context is provided by a set of generative mechanisms that connect the different levels lexical semantic representations. These generative operations include *type coercion*, *selective binding* and *co-composition*, and are viewed as semantic transformations operating over type combinations in grammar to facilitate well-formedness.

¹³ A more detailed discussion of this issue follows in section 2.3.2.3 below.

2.3.2.1 TYPE COERCION

The mechanism of *type shifting* has been developed largely in works by Rooth and Partee (1982) and Partee and Rooth (1983) to explain how expressions in language may change their type depending on the context. In addition, a *type ladder* has been introduced to capture the relations between the different types of an expression.

TYPE COERCION is a lexically governed type shifting which is defined by Pustejovsky (1995, p 111) as *a semantic operation that converts an argument to the type which is expected by a function, where it would otherwise result as a type error*. Therefore, this generative device accounts for the cases where a lexical item or phrase is forced by the governing item in the phrase to accommodate a semantic interpretation without a change of its syntactic type.

The application of *the type coercion* mechanism allows a proper treatment of verbs which take different complement types as the examples in (25).

- (25) a. Susan wants *a martini*. (NP)
 b. Susan wants *to go*. (VP [+INF])
 c. Susan wants *Alex to go*. (S [+INF])

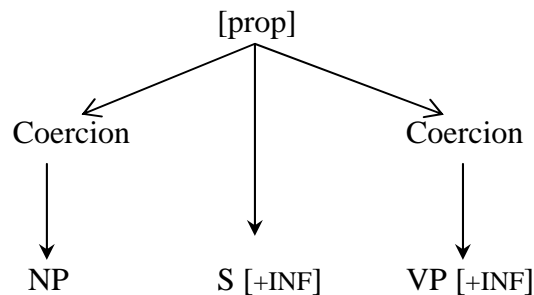
The traditional approach to verbs of this kind (cf. Dowty 1979, 1985, among others) is to regard the verb at hand as ambiguous and consider it as several different lexical entries. Such sense enumeration approaches would ascribe different meanings to each of the possible uses of *want* in the various subcategorization patterns. Thus, the sentences in (25) would be considered as examples of three different lexical entries as seen in (26) below.¹⁴

- (26) a. $want_1 \in \langle NP, NP, S \rangle$
 b. $want_2 \in \langle VP, \langle NP, S \rangle \rangle$
 c. $want_3 \in \langle S, \langle NP, S \rangle \rangle$

¹⁴ The meaning postulates in (26) are suggested by Dowty (1985) and also repeated in Pustejovsky (1995).

Instead of postulating different meanings of the verb for each subcategorization frame so that it would fit in the various syntactic environments, Pustejovsky (1995) proposes a *type coercion* that is lexically governed by the verb and converts the complement to the type which is expected by the verb. Thus, although *want* may take complements which have different syntactic expressions it is the semantic type that is taken into account instead of the possible syntactic forms in which it may be realized. This is exemplified in Pustejovsky (1995) by the tree replicated in (27) below, where the semantic type of the complement of *want* relates to different syntactic expressions.

(27)



2.3.2.2 SELECTIVE BINDING

SELECTIVE BINDING obtains when *a lexical item or phrase operates specifically on the substructure of a phrase, without changing the overall type in the composition* (Pustejovsky 1995, p. 61). This semantic mechanism treats the lexical item as a function and applies it to a particular quale of the other item in the composition. Thus, a selective interpretation is achieved.

An illustration of the application of the selective binding rule with respect to the problem of adjectival polysemy is presented through the examples in (28) and the subsequent analyses following Pustejovsky (1995).

- (28) a. My friend is a fast typist.
 b. To cut this meat I need a good knife.

c. This is a good knife but it doesn't cut very well.

The interpretation of *fast typist* in (28a) cannot be derived by the formula $\lambda x[\textit{typist}'(x) \wedge \textit{fast}'(x)]$, which will result in the rather different description of a person who is a typist and a person who is fast. Instead it is suggested that the adjective is selecting a particular quale within the head noun of the phrase it participates in. Thus, for the noun *typist*, which will have a representation as the one in (29) below,

$$(29) \left[\begin{array}{l} \textit{typist} \\ \text{ARGSTR} = [\text{ARG}_1 = x : \textit{human}] \\ \text{QUALIA} = \left[\begin{array}{l} \text{FORMAL} = x \\ \text{TELIC} = \textit{type}(e, x) \end{array} \right] \end{array} \right]$$

the proper interpretation will have the following format as in (30).

$$(30) \lambda x[\dots \textit{Telic} = \lambda e[\textit{type}'(e, x) \wedge \textit{fast}'(e)] \dots]$$

In addition, the examples in (28b) and (28c) show that the adjective doesn't select for qualia, but for a specific type within the quale. Thus, the formal representation of a *good knife* as used in (28b) will refer to the type specified in the TELIC quale as shown in (31).

$$(31) \lambda x[\dots \textit{Telic} = \lambda e[\textit{cut}'(e, x, y) \wedge \textit{good}'(e)] \dots]$$

This will result in the interpretation of a *good knife* as 'a knife that cuts well' in (28b), as opposed to the interpretation in (28c) where *good* will refer to the knife as an artefact and will instead denote 'a knife that is well made,' thus referring to the AGENTIVE quale of *knife*, as can be seen in (32) below.

$$(32) \lambda x[\dots \textit{Agentive} = \lambda e[\textit{create}'(e, y, x) \wedge \textit{good}'(e)] \dots]$$

2.3.2.3 CO-COMPOSITION

CO-COMPOSITION takes place when *multiple elements within a phrase behave as functors, generating new non-lexicalized sense for the words in the composition* (Pustejovsky 1995, p. 61). Cases of underspecified semantic forms that may become contextually enriched, such as *manner co-composition*, *feature transcription*, and *light verb specification*, are also seen as applications of exactly this generative mechanism.

One of the cases of verbal logical polysemy mentioned above, namely the case of *bake*, provides a good illustrative example of how the appliance of co-composition will account for the generativity in semantics, thus avoiding lexical enumeration.

In Atkins et al. (1988), the verb *bake* is considered to exhibit two distinct meanings, namely a *change of state* and a *creative event*, as shown in the examples in (33a) and (33b) respectively.

- (33) a. My brother baked the potatoes. (*change of state*)
 b. I baked the cake. (*creative event*)

Consequently, in Levin (1993) and Levin and Rappaport (1995) these senses are listed as separate lexical entries, which has already been discussed as redundant.

As previously described in chapter 2.3.1.3, the AGENTIVE quale of a lexical item specifies the factors involved in its origin. Hence the AGENTIVE value of an artefact, like *cake*, will refer to the act of creating the object denoted by the particular word; in this case it will be the process of *baking the cake*. Thus, a referential relation is established between the value of the AGENTIVE quale of *cake* and the process which is denoted by the verb *bake*, and which is also specified in the AGENTIVE quale of the verb. This relation, dubbed *co-specification* (Pustejovsky 1991), results in a type feature unification of the AGENTIVE values of the verb and its argument, thus licensing a semantic operation named *qualia unification* (Pustejovsky 1995).

Therefore, the verb *bake* is not polysemous. Instead, its *creative event* sense is derived when a complement, which co-specifies the verb, extends its basic meaning by co-composition.

Similar interpretation is assumed for the *Resultative* construction and the cases of directional phrases with non directional verbs.¹⁵ Thus, new senses of the words and verbs in particular, are not simply posited and hence enumerated in the lexicon. Instead, they are generated by means of semantic operations, which follow naturally from the interaction between words in a sentence.

In my own research, these semantic operations proved to be a useful tool in analysing cases where verbs were used with extended meaning to denote situations that were not encoded originally in their lexical representation. More detailed analyses and discussions of such cases follow in Chapter 4 (section 4.3) and Chapter 5 (section 5.2) of the present work.

2.4 Conceptual Semantics and the multi-tiered representation of CS

Conceptual Semantics evolved in the early 80s (Jackendoff 1983) as a combinatorial system of considerable complexity and subtlety, very similar to the ones already proposed for syntax (Chomsky 1965) and phonology (Goldsmith 1976). Conceptual Semantics is concerned most directly with the organization of the internal mental representations that constitute the conceptual structure, and with the formal relations between this and other levels of representation.

Conceptual structure (CS), as introduced in Jackendoff (1983), is the form in which speakers encode their construal of the world. Lexical concepts are the concepts expressed by the words in the sentence and are therefore the basic units out of which a sentential concept is construed. Learning a lexical concept, then, is construing a composite expression within the grammar of lexical concepts and associating it with phonological and syntactic structures, and storing them together in the long-term memory as a usable unit for future access. The grammar of lexical concepts consists of a finite group of mental primitives and principles of mental combination that collectively determine the set of lexical concepts. The theory of conceptual semantics thus takes conceptual structure to be rather parallel to the syntactic and phonological structures.

¹⁵ These are discussed in Levin and Rapoport (1988) as cases of lexical subordination.

2.4.1 *Conceptual Semantics*

The theory of Conceptual Semantics, as further developed in Jackendoff (1990), pays special attention to the lexicon and its entries and explores the composite nature of conceptual structures. An organisation of the mental information structure involved in language is introduced to include three distinct autonomous levels of structure: phonological, syntactic and conceptual. Each of these levels has its own organisation into subcomponents, its own set of primitives and principles of combination, and is described by a set of formation rules that generates the well-formed structures of the level. There are also sets of correspondence rules that link the levels and rules of inference for the conceptual structure domain.

In addition, correspondence rules between the linguistic levels and non-linguistic domains are also included. These rules determine the mapping from the auditory input into phonological structure, as well as the mapping from phonological structure into the motor output. On the level of conceptual structure, however, the correspondence rules determine the mapping between conceptual structure and other levels of mental representation that encode the input and output of vision, action, etc.

This conception of mental organisation or the so-called Representational Modularity is explored in Jackendoff (1997). The main idea of the modular view of language and cognitive capacities is that there is some finite number of distinct modules of the mind and each of these modules is responsible for a different representational format or *a language of the mind*. Each of these *languages* is a formal system with its own set of primitives and principles of combination. Following this idea Jackendoff proposes a tripartite parallel generative architecture where both phonology and semantics generate structures together with syntax.

However, he argues that only phonology and syntax are proper language systems while concepts are part of the lexicon but not of language itself, since language is not necessary for the use of conceptual structure. There are possible situations where non-linguistic organisms (primates or babies, for example) use conceptual structures as part of the encoding of their understanding of the world. Jackendoff sees Conceptual Structure as much richer and including other types of thought. He also posits a system

of interface modules in addition to the representational modules. Thus, an interface between system A and system B is seen as consisting of three components as in (34).

- (34) a. A set of representations in system A to which the interface has access,
 b. A set of representations in system B to which the interface has access,
 c. A set of A-to-B correspondence rules.

The correspondence rules, however, do not perform derivations in the standard sense of mapping a structure within a given format into another structure within the same format. Crucially, they map one representational format into another, as for example, phonetic representations into motor instructions during speech production.

Thus, the organization of the Grammar as a mental information structure and the relations of the module of Conceptual structures to the rest of the representational and interface modules is diagrammed in Fig.7 below.¹⁶

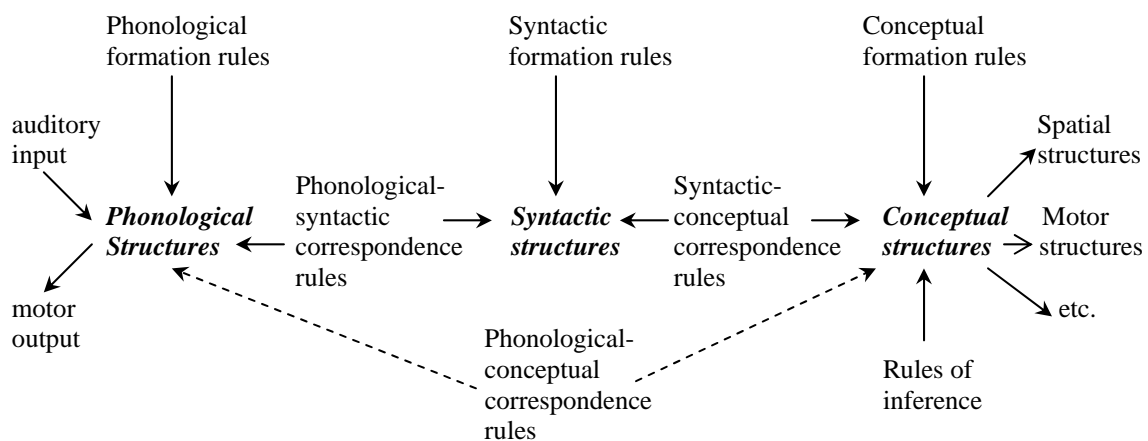


Fig.7 The organization of grammar within the cognitive architecture of human mind

¹⁶ This diagram is a blend of the diagrams presented in Jackendoff (1987), Jackendoff (1990), and Zee & Nikanne (2000) thus aiming at a richer representation of the organization of Grammar within the the cognitive architecture of human mind, as originally proposed by Jackendoff, and further elaborated by Nikanne and Zee.

The apparent missing component in this figure is the lexicon, which according to Jackendoff (1990) cuts across the other modules, as each of them is additionally divided into lexical versus extralexical principles. This is visually presented in the schematic representation of the linguistic system given in Nikanne (1996) where a formal insertion of (*the lexicon*) component is introduced as being connected to all the representational modules via linking rules.

With respect to the strong association between lexical items and the conceptual structure representation, a more detailed picture of the status of lexical item as discussed by Jackendoff (2002) is presented in the next section.

2.4.2 Conceptual Structure and the lexical items

Conceptual structure, as already discussed in the previous section and visually presented in Fig.7 above, is a central cognitive level of representation, which links the linguistic modules of phonology and syntax with other cognitive capacities such as auditory, visual, spatial, etc. In this view a rethinking of the term *lexical item* was necessary and in a recent paper titled "*What's in the Lexicon*" Jackendoff (2002) addresses precisely these issues.

Until recently, a widespread stereotype and a popular conception of language was that the memorised units of language (and therefore the ones stored in the lexicon) were *words*. Thus, the terms *lexical item* and *word* were used interchangeably, and the term *lexicon* was to stand for all the words the speaker knows and therefore contained only non-predictable features. As it turned out, however, these assumptions deviated from the psychological reality.

Therefore, Jackendoff (2002) argues that lexical items may also be smaller or bigger than grammatical words, that not all grammatical words are lexical items, and that there are lexical items which contain no phonological form. A lexical item is seen, then, as a tripartite multiple interface rule, where the three components phonology, syntax and semantics can exist independently of each other. This can be observed in the

so-called *defective words* (ibid, p. 27), where one of the components may not be present as illustrated in the examples in (35) below.

- (35) a. interjections (*ouch, hello...*) – phonology and semantics, no syntax
 b. *do*-support, expletives (*it, there*) – phonology and syntax, no semantics
 c. PRO (subject of infinitive as in *He tried (PRO to leave)*) – syntax and semantics, no phonology

A possible solution is found in a heterogeneous theory of the composition of lexicon distinguishing between productive (for example, derivation, inflection) and semi-productive (irregular verbs) processes, depending on where morphology is, in lexicon or grammar, or in both. Then following Jackendoff's proposal, *girl* is both a word and a lexical item; *-s* is a lexical item but not a word; *girls* is a word but not a lexical item, since it is constructed from two lexical items *girl* and *-s* ; and the irregular form *women* is both a word and a lexical item. How is this information mentally represented?

Earlier theories have speculated that this may be a single defining feature of the item or sets of necessary and sufficient conditions. However, it is most likely that mental representations are lists of linguistic features, which in addition may include input from other modalities, as proposed in Jackendoff (2002). Thus, the mental representation of the word *cat*, for example, could include the information given in (36) below.

(36) <i>cat</i>	<table border="0"> <tr> <td>Phonology</td> <td>/k æt/</td> </tr> <tr> <td>Syntax</td> <td>Noun</td> </tr> <tr> <td>Conceptual Structure</td> <td>Thing</td> </tr> <tr> <td>3 - D model</td> <td>(picture)</td> </tr> <tr> <td>Audial representation</td> <td>(sound)</td> </tr> <tr> <td>Haptic representation</td> <td>(warm, soft fur ...)</td> </tr> <tr> <td>Etc.</td> <td>...</td> </tr> </table>	Phonology	/k æt/	Syntax	Noun	Conceptual Structure	Thing	3 - D model	(picture)	Audial representation	(sound)	Haptic representation	(warm, soft fur ...)	Etc.	...
Phonology	/k æt/														
Syntax	Noun														
Conceptual Structure	Thing														
3 - D model	(picture)														
Audial representation	(sound)														
Haptic representation	(warm, soft fur ...)														
Etc.	...														

The crucial points in this lexicalist view are that no modular dissociation between grammar and lexicon is assumed and there is also a reduction of the traditional distance between linguistic theory and psychological and developmental considerations. In addition, the answer to the question of what is stored in the mental lexicon and what is computed online has also a significant impact on the organization of the theory of grammar. Thus, Jackendoff (2002) suggests a new approach where the theory of competence must be seen as dynamically interacting with the theory of performance. Accounting for this fact would facilitate linguistic exploration of the most suitable representational format of lexical items, which has been a central consideration in my research, too.

2.4.3 The tiered representation of Conceptual Structure (CS)

The Conceptual Structure representation (Jackendoff, 1987) is motivated by several basic principles within the theory of Conceptual Semantics¹⁷ – the requirement for finite representability, the applicability of concepts in unique situations, and the learnability of concepts from *a sufficiently rich innate basis* (ibid. p.375).

The organization of Conceptual Structure, as introduced by Jackendoff (1987, 1990)¹⁸ and further developed by Nikanne (1990, 1995), involves at least three separate tiers: (i) the thematic tier, (ii) the action tier, and (iii) the temporal tier. These are conceived as independent combinatorial systems that are interconnected by sets of interface rules. The purpose of these tiers is to handle different dimensions of the representation of Conceptual Structure, thus accounting for the different roles of the participants in the situation denoted by the verb at hand.

¹⁷ These principles are regarded as parallel to the principles in generative syntax and phonology (Jackendoff, 1987).

¹⁸ Throughout the development of the theory, the different elements of the tiers have changed names. However, the organization of Conceptual Structure as a whole remains relatively steady.

2.4.3.1 The thematic tier

The thematic tier is regarded as the earliest layer of conceptual structure, as it was formally introduced in Jackendoff (1983, 1987), and according to Nikanne (1990) an elementary structure of the thematic tier can be found already in Jackendoff (1972).

The thematic tier consists of a set of primitive conceptual categories such as THING, EVENT, STATE, ACTION, PLACE, PATH, PROPERTY, and AMOUNT. In addition, formation rules operate on these categories to generate more complex structures through a set of functions as illustrated in (37) below.

$$(37) \quad \begin{array}{l} \text{a. } [\text{EVENT}] \rightarrow \left\{ \begin{array}{l} [\text{Event GO}([\text{THING}], [\text{PATH}])] \\ [\text{Event STAY}([\text{THING}], [\text{PLACE}])] \\ [\text{Event MOVE}([\text{THING}])] \\ [\text{Event INCH}([\text{STATE}])] \end{array} \right\} \\ \\ \text{b. } [\text{EVENT}] \rightarrow [\text{Event CAUSE} ([\left\{ \begin{array}{l} \text{THING} \\ \text{EVENT} \end{array} \right\}], [\text{EVENT}])] \\ \\ \text{c. } [\text{PATH}] \rightarrow \left[\begin{array}{l} \left\{ \begin{array}{l} \text{TO} \\ \text{FROM} \\ \text{TOWARD} \\ \text{AWAY - FROM} \\ \text{VIA} \end{array} \right\} \left(\left(\left[\begin{array}{l} \text{THING} \\ \text{PLACE} \end{array} \right] \right) \right) \end{array} \right] \\ \text{Path} \end{array}$$

As mentioned above, the thematic tier includes also a number of functions which in Nikanne (1990, 1998) are additionally divided into three zones, on the basis of common features shared among the various functions within each zone. Every zone covers a different thematic role type. A schematic representation of the zones with the distribution of the various functions of the thematic tier and the respective thematic roles of the three zones is offered in Fig.8 below.¹⁹

¹⁹ Fig. 8 presents the distribution of the zones within the thematic tier as given in Nikanne (1998).

	(3) Causative Zone	(2) "Figure Zone"	(1) "Ground Zone"
non-monadic functions:	CAUSE LET	GO, BE, STAY, ORIENT, EXTEND	
monadic functions:	INCH(oative)	MOVE CONFIG	TO, FROM, VIA, TOWARD, AWAY-FROM; AT, IN, ON, UNDER, NEAR ...
thematic role:	Agent (= "causer")	Theme (= "Figure")	Location (= "Ground")

Fig. 8 Zones within the thematic tier

Besides the division of the functions into the three zones, there is also a certain alignment across the zones. Thus, the functions are again grouped into monadic and non-monadic. As the names suggest, monadic functions can have only one complement, which can be either an argument or another function, while non-monadic function allow for more than one complement.

The terms *Agent* and *Theme* are not regarded as semantic primitives (Jackendoff, 1987). Instead they are considered as relational notions which can be defined structurally on the basis of conceptual structure. Thus, their status is viewed as parallel to that of Subject and Object in syntactic theories like the Government and Binding approach, to mention some.

With respect to the formation rules in which the various functions operate, we can distinguish several types of functions as PATH-functions, STATE-functions, EVENT-functions, and PLACE-functions. Thus, for example, the CAUSE function²⁰ is an EVENT-function, whose first argument is Agent and second argument is Event, caused by that Agent. The successful outcome of the function is marked by the presence of a value which can be assigned to it with one of the three superscript symbols (+), (-), or (u). Thus, CS⁺ represents a positive outcome of the function, CS⁻ stands for a negative

²⁰ Jackendoff (1990) uses also CS to denote the more abstract function; thus CAUSE was used initially to notate standard causation, which after the introduction of the superscripts is notated as CS⁺.

outcome, and CS^u is used when the outcome of the CAUSE function cannot be determined. These three cases are illustrated in (38a), (38b) and (38c) respectively with sentences similar to those used by Jackendoff (1990).

- (38) a. Emily managed to run away. (CS⁺)
 b. Emily failed to run away. (CS⁻)
 c. Emily tried to run away. (CS^u)

In addition, the functions are related by a set of principles which govern the successful embedding of different functions in a *function chain (f-chain)* (Nikanne 1995). The f-chain is regarded as a headed structure (Nikanne 1990, 1995) in which the scope of the *head-complement* relations goes from left to right, and also from the f-chain to the theta arguments. Thus, first the f-chain assigns a thematic role to each argument and then the arguments are assigned act-roles in the action tier.

2.4.3.2 The action tier

The action tier, introduced in Jackendoff (1987, 1990) and explored in Nikanne (1995), can be briefly presented as consisting of two dyadic functions AFF and REACT which map their arguments into a State or an Event. These functions are considered the primitives of this tier. The purpose of the action tier is to express dominance relation between the two basic participants Actor and Undergoer.

The AFF function operates over two arguments, the first of which is Actor and dominates over the second, Undergoer. The argument Undergoer subsumes the traditionally divided roles of Patient and Beneficiary, while the same distinction is achieved with assigning different values to the function AFF. Thus [AFF⁺ ([X],[Y])] represents a situation where X affects positively Y (i.e Y is Beneficiary) like in '*Mary assisted John.*' [AFF⁻ ([X],[Y])] stands for a situation where X affects Y in a negative way (i.e. Y is regarded as Patient) like in '*Nick stabbed Tom.*' And [AFF^u ([X],[Y])] symbolizes a situation where the affect cannot be judged as positive or negative and Y is thus 'neutral undergoer' as in '*Susan drove her daughter to school.*' Besides, AFF⁰ is

used to symbolize a neutral reaction or non-opposition in account for the relation of letting as in '*Susan allowed her to stay.*'

In addition, the features [\pm *volitional*] are introduced to distinguish between a deliberate ([+vol]) or an accidental ([-vol]) involvement of an animate Actor as illustrated by Jackendoff (1990) in an example similar to the one given in (39) below.

(39) The boy rolled down the hill.

$$\left[\begin{array}{l} \text{GO} ([\text{BOY}], [\text{DOWN} [\text{HILL}]]) \\ \left\{ \begin{array}{l} \text{a. AFF}^+_{\text{vol}} ([\text{BOY}], \quad) \\ \text{b. AFF}^-_{\text{vol}} ([\text{BOY}], \quad) \\ \text{c. AFF} \quad (\quad, [\text{BOY}]) \end{array} \right\} \end{array} \right]$$

The sentence in (39) can be associated with three different action tiers depending on the presence and the value of the *volitional* feature. Thus, the structure in a) represents a wilful doer, the one in b) represents a nonwilful doer, while c) stands for an undergoer.

The REACT function introduced in Jackendoff (1990) is described by Nikanne (1990) as opposite to AFF in a sense that the first argument is *reacting to the effect of the second argument*. Similarly to AFF, REACT can also appear with three different values which symbolize the character of the reaction of the first argument as result of the effect of the second argument. The three possibilities positive (+), negative (-), or neutral (u) are illustrated in the examples in (40) below.

- (40) a. [REACT⁺ ([X],[Y])] as in *Mary enjoyed her dinner.*
 b. [REACT⁻ ([X],[Y])] as in *Mary disapproved his behaviour.*
 c. [REACT^u ([X],[Y])] as in *Mary considered him funny.*

Likewise the f-chain in the thematic tier, a sequence of successfully embedded action functions is called *act-chain* and again there are a number of principles that allow for well-formed chains while ruling out defective structures.

Thus, the conceptual representation of a sentence like the one in (41) will combine the two tiers presented so far, namely the thematic tier and the action tier, and will have the structure in (42) below.²¹

(41) The boy cast a stone into the river.

(42)	AC -----UN		<i>act-chain</i>
	BOY	STONE	arguments
	CAUSE =====	GO === TO == IN	<i>f-chain</i>

As argued in Nikanne (1998) the linking between the two chains is based on free principles, whereby the act-chain selects its arguments from the f-chain.

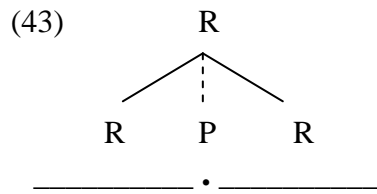
The alignment of elements from the different tiers is very similar in function to the set of features ascribed to a participant at the different dimensions in the Sign Model (Dimitrova-Vulchanova 1996/99, Hellan & Dimitrova-Vulchanova 2000) to be discussed in Chapter 3, section 3.3.

2.4.3.3 The temporal tier

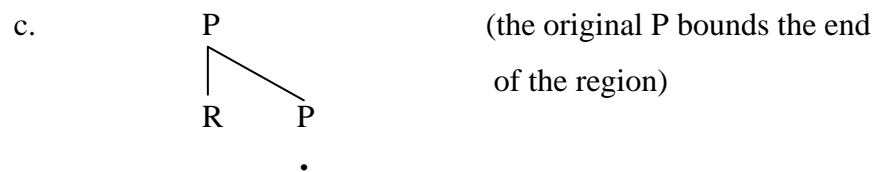
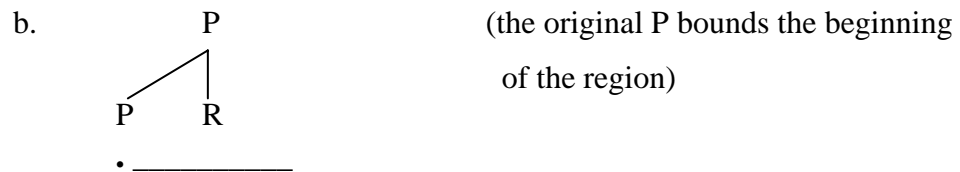
In addition to the two main tiers presented above, Jackendoff (1987) proposes the existence of a temporal tier which consists of two primitives: point of time (P) and region of time (R). While R is one-dimensional entity and can be presented as a left-to-right directed line, P is considered a zero-dimensional entity and can be illustrated as a point.

The organization of the temporal tier is further structured (Jackendoff 1990, also mentioned in Nikanne 1990) as consisting of two main possibilities. One can either concentrate on a period of time and discover that it includes a point as presented in (43) below,

²¹ The sentence in (40) is similar to some of the examples presented in Nikanne (1990, 1995, 1998).



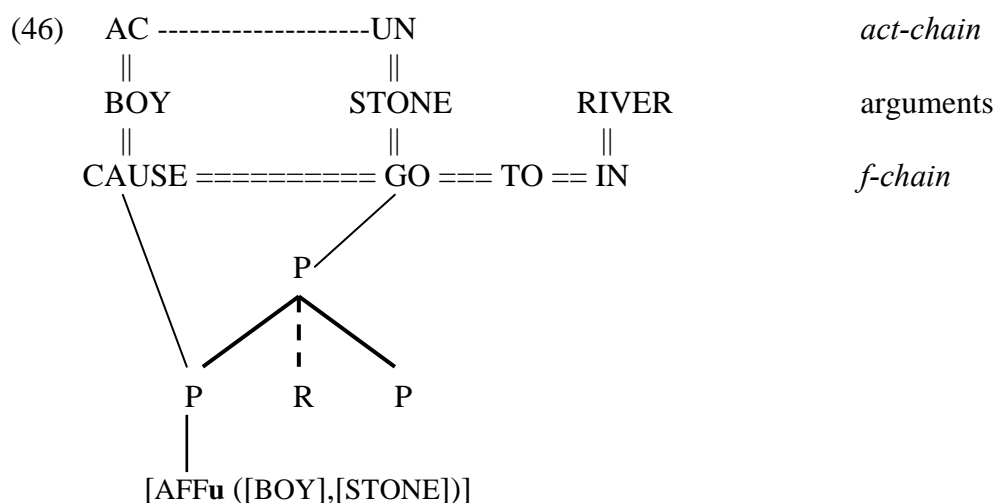
or one can examine a point of time and discover that the internal structure of the point includes a region, too. In the latter case, there are three options, presented in (44a) to (44c)²² respectively.



The elements of the temporal tier are related to the functions of the thematic tier and the action tier as in the example in (45) below and subsequent representation in (46).

(45) The boy cast a stone into the river.

²² The schematic representations in (40a) to (40c) are adopted from Nikanne (1990).



In the representation in (46) above the thin lines signify the connections with the functions in the other tiers, while the thick lines symbolize the organization within the temporal tier.

2.5 Conclusions

In this chapter I have presented some of the current approaches to the Syntax-Semantics Interface in their relation to my research.

First, I discussed Levin's verb classes and alternations and while I agreed with her on the hypothesis that the meaning of the verb determines to a large extent the verb's subcategorization preferences, I stand apart on the assumption that verbs can be classified straightforwardly into classes, without paying attention to the subtlety in their meaning and the differences in the events they can lexicalize.

Next, I have presented the theory of Construction Grammar since it is assumed to explore verb meaning in relation with sentential meaning. While the Construction Grammar approach might be successful in explaining the behaviour of certain grammatical constructions, it contributes little to the general understanding of lexical meaning. Therefore, I opposed the idea of postulating constructions as completely detached meaningful units. Instead, I have suggested an explanation on which different types of situations should be distinguished, thus accounting for the possibility of

describing the same situation type in similar way, hence resulting in construction-like patterns, which do not carry meaning on their own, while still allowing different verbs to behave in the same way, when lexicalizing an event of particular type.

In section 2.3 I have brought forward some of the main ideas of the Generative Lexicon framework as it focuses on the generativity of word senses in natural language, thus explaining the interpretation of words in context and accounting for systematic relatedness between word senses in a formal and predictable way. I particularly consent to the view that there exist a number of generative mechanisms in semantics, which operate over lexical items in composition, thus generating new senses that do not need to be enumerated as separate lexical items. In addition, an extended event structure is posited, which embodies a fine-grained representation of the internal configurational properties of the event. This is in line with recent research on verb semantics (cf. Dimitrova-Vulchanova 1996/99, 2003, Hellan & Dimitrova-Vulchanova 2000), as well as with my own research discussed in Chapters 4 and 5.

Finally, I have presented the theory of Conceptual Semantics and the tiered representation of Conceptual Structure as it is very close in agenda to the Sign Model framework which has been used as the main model of analysis in my work and thus served as a natural link to the next chapter.

3. Theoretical Background

My research has been particularly influenced by several theoretical frameworks and experimental studies. These have served as basic guidelines for the collection and the analyses of the empirical data used in the research.

The theoretical approaches to be discussed in this chapter stand very close in their concern for the syntax-semantics interface and particularly in exploring what items from Conceptual structure can be assumed to be present in the semantic representation of lexical units based on criteria other than obligatory syntactic realization. Thus, they were helpful devices towards finding a proper representation, which will truthfully reflect the semantic properties of lexical items and their mapping on syntax.

Section 3.1 presents a relevant theoretical approach by Koenig, Mauner, & Bienvenue (2002, 2003) on lexically encoded participant information, together with the experimental studies conducted to test this hypothesis.

Then, in section 3.2 I outline the theoretical basis of the proposal by Hare, McRae & Elman (2003) that verb subcategorization preferences are contingent on verb sense. Then I briefly present the results of their experiments carried out to test the correlation between verb sense and the verb's preferred subcategorization frames.

Finally, section 3.3 brings forward the main ideas of a formal framework currently known as *The Sign Model* (Dimitrova-Vulchanova 1996/99, Hellan & Dimitrova-Vulchanova 2000) adopted here as the primary model for the formal analyses of the empirical data.

3.1 Lexical Encoding of Participant Information

In their paper "*Arguments for Adjuncts*," Koenig, Mauner, & Bienvenue (2001¹, 2003) argue that although it is widely accepted that the syntactic structure of many sentences is determined mostly or entirely by the participant information included in the lexical entries of verbs, there are no reliable syntactic criteria that can be used to delimit the set of items that can express lexically encoded participant information. In other words, there has not been established a set of necessary and sufficient syntactic criteria that can serve as a clear-cut basis for the distinction between information that is *lexically encoded* and information that is not (that is the argument/adjunct distinction).

In what follows (sections 3.1.1 through 3.1.3), I will briefly describe the notion of *lexically encoded participant information*, as defined by Koenig et al. (2002, 2003) and I will present some of the most common syntactic criteria, which however are not sufficient for the successful authentication of the information that is encoded in the lexical semantic representation of verbs. In section 3.1.3 I will argue in line with Koenig et al. (2002, 2003) that a set of semantic criteria is necessary to account for the argument/adjunct distinction. Section 3.1.4 describes the studies conducted by the authors to test the semantic criteria proposed.

3.1.1 *Lexically encoded vs. non-lexically encoded participant information*

It is generally assumed across current linguistic theories that there is an important dichotomy between the different syntactic constituents of the sentence, and that some of them are required by and dependent on the verb at hand, while others may co-occur with various verbs, and are not necessary. Thus, the participants whose presence is directly connected to the main verb in the sentence are usually called *arguments* or *complements* of the verb, while the others are called *adjuncts* or *modifiers*. To avoid terminological

¹ Besides the revised version published in 2003, there is also an earlier more detailed version of the paper which was available on-line since 2001 and is therefore included in my bibliography. For convenience, I will only refer to the year of the published edition (2003) in subsequent inline citations.

confusion, Koenig et al. apply the term *arguments* to the lexically encoded semantic participants and *adjuncts* to the non-lexically encoded semantic participants only. The terms *complements* and *modifiers*, respectively, are used to refer to the syntactically co-occurring constituents that are lexically encoded or not.

Lexically encoded participant information is defined as the information about the participants which take part in the situation denoted by the main verb, as well as about the nature of their participation. Namely this is the information which must be included in the semantic representation of the verb at hand. This information, however, may be given different representational formats. On the account of Koenig et al. (2002, 2003), the lexical encoding of participant information reduces to semantic categories which are activated upon recognition of a word, as they assume a mental lexicon which can be described as a multidimensional hierarchy of categories, which combine syntactic, semantic and morphological information (e.g. along the lines of the HPSG framework).

The information that is neither required nor dependant on the particular verb is called *non-lexically encoded participant information*. Thus, regardless of whether it is syntactically present or not, non-lexically encoded information is not semantically compulsory to identify the situation denoted by the verb at hand. Thus, for Koenig et al. lexically encoded information is exclusively defined semantically and psychologically, which is also supported in their experimental data.

However, in the common tradition outside their work, there has not been established a clear-cut distinction between lexically encoded and non-lexically encoded participant information. Thus, in the example in (1) below, it was generally assumed based on syntactic criteria that only the NPs *the boy* and *his friend* are true arguments of the verb, while all the PPs are considered as adjuncts and hence excluded from the lexical representation of the verb.

- (1) The boy hit his friend [*on the head*] [*with a stone*] [*yesterday*] [*in the schoolyard*].

In line with Koenig et al. many of the recent theoretical approaches (cf. for example Pustejovsky 1995, Dimitrova-Vulchanova 1996/99), as well as independent research on

empirical data from a variety of languages (Donohue C. & M. Donohue 2004, to mention some) suggest that this is not the case.

3.1.2 Possible syntactic criteria and their inadequateness

Various syntactic diagnostics have been suggested as tests for argumenthood.² Although some of them may seem as concrete criteria for the argument/adjunct distinction, they all presuppose the knowledge of the existence of such a division at first place, thus they cannot serve as the basis for the distinction between arguments and adjuncts itself.

However, some of the proposed syntactic criteria do have semantic correspondences which are reliable in that they can be independently observed by language users, and precisely these are presented below in sections 3.1.2.1 to 3.1.2.3.

3.1.2.1 Syntactic obligatoriness

Traditionally, one of the strongest criteria for argumenthood was considered to be the syntactic obligatoriness of a participant. If a participant cannot be omitted without this resulting in the ungrammaticality of the sentence, it is defined as an *argument*. Thus, *the boy* and *his friend* in the example in (1) are predicted to be the only syntactic arguments of the verb *hit*.

Although all syntactically obligatory participants are in fact semantic arguments, the opposite implication is not correct. That is, not all syntactically optional constituents should be directly identified as adjuncts, as it can be seen in the example in (2).

- (2) a. Susan sent him a letter.
b. Susan sent a letter.
c. Susan sent a letter to her friend Peter.

² Most of these tests have been summarized and discussed in Schütze (1995) among others.

The verb *send* is commonly assumed to be ditransitive, thus subcategorizing for both *him* and *a letter* in the sentence in (2a). However, the italicized phrase in (2c) is not obligatory for the grammaticality of the sentence as seen if compared with (2b). Thus, the argument, usually identified as *recipient* or *goal* (in this case *him* in (2a) and *to her friend Peter* in (2c)) is not syntactically obligatory, although it is semantically present.

Koenig et al. (2003) suggest that syntactic obligatoriness is not only an insufficient criterion but also that it cannot constitute even a necessary condition for a participant to be judged as an argument. Data from Spanish reflexive passives show that the agent phrase³ cannot be overtly expressed as we can see in the example in (3) below.⁴

- (3) a. Se cometieron varios atentados (*por los terroristas) para amedrentar a la policía.
 refl.committed-3p.pl several attacks (by the terrorists) for intimidate-INF the police
 a' Several attacks were carried out (by the terrorists) to intimidate the police.

The existence of an agent in the situation denoted by the main verb in the example in (3) is suggested by the grammaticality of the purpose clause *para amedrentar a la policía* (to intimidate the police). It has also been argued (cf. Dimitrova-Vulchanova 1995, 1996/99) that the teleology of passives is the demotion of the agentive participant. Therefore, such a participant must have been present in the situation in the first place. Nevertheless an overt syntactic expression of the agent in Spanish⁵ is not possible as seen in the ungrammaticality of the phrase *por los terroristas* (by the terrorists) in the Spanish sentence in (3a) as opposed to its English translation in (3a').

³ As noted by Keenan (1985) the term *agent phrase* can be misleading as its real thematic role depends on the verb of which it is the understood subject and does not have to be exclusively Agent.

⁴ Thanks to Liliana Serbezova-Martinez for this example.

⁵ As suggested by Koenig et al. (2003) following Keenan (1985), the phenomenon observed in Spanish is attested cross-linguistically and thus may serve as a basis for generalization.

Thus, a participant need not be syntactically expressed in order to be conceptualized by language users as taking part in the situation denoted by the verb at hand, and therefore be perceived as encoded in the semantic representation of that verb. Therefore, syntactic obligatoriness cannot be considered as neither a sufficient nor a necessary criterion for a participant to be judged as present or absent from the lexical representation of verbs.

3.1.2.2 Iteration of participants

A commonly used test for adjuncthood (cf. Fillmore 1968, Pollard and Sag 1987, and Schütze 1995, among others) is the possible iteration of adjuncts within the simple sentence as opposed to the impossibility of arguments to be iterated. This contrast is often illustrated with examples like those in (4a) and (4b) below.

- (4) a. *Mary gave the book *to John, to Bill*.
 b. Mary met them *in the church, in Oslo, at 10 a.m., on Sunday morning*.

Thus, it is assumed that *goals* or *recipients* cannot be iterated as shown in (4a), while *locations* and *times* can, and should therefore be considered as adjuncts. However, the examples in (5) below suggest that this cannot be the case.

- (5) a. They stabbed him *in the leg above the knee*.
 b. *They stabbed him *in the leg in the foot*.
 c. *She met them *in the school, in the church*.⁶

The example in (5a) demonstrates that *goals* can in fact be iterated, as *in the leg* would traditionally be defined as *goal*, while the example in (5c) shows that locations may have certain restrictions on iteration. Then what is the difference in the ungrammaticality of (5a) and (4b) as opposed to (5b) and (5c) respectively?

⁶ On the interpretation that the school is not inside the church.

Koenig et al. (2003) suggest that iterated locations and times must describe a progressively wider region or space, thus jointly specifying *a single participant at two different levels of granularity*.

Consider the examples in (6) and (7) below.

- (6) a. He put the pictures *(in the drawer in the common room).
 b. She lives/resides *(in New York, in a big red house).
 c. I left the keys in the house (on the round table).
- (7) a. *I ate the meal with a spoon, with a fork.
 b. *Mary saw them on Friday, on Sunday morning.
 c. *He talked happily, gladly/miserably.
 d. He talked fast, happily.

The examples in (6a) through (6c) show that constituents which are syntactically obligatory and traditionally considered as arguments *can be iterated*; and in the case of (6b) and (6c) the location phrases do not subsequently define a wider region, but a narrower one.

In contrast, the examples in (7a) to (7c) demonstrate that constituents which are commonly considered as adjuncts *cannot be iterated* if this violates the semantic continuum of the event and results in a semantic contradiction.

Thus, I will suggest that neither arguments nor adjuncts can be freely iterated unless they specify a single participant at different levels of granularity (which need not be progressively wider), or in different dimensions, as seen from the contrast of (7c) and (7d). That is, the information provided by the iterating constituents should not be redundant or controversial in any way. Therefore, iteration cannot be used as a reliable diagnostics for adjuncthood and no participant can be judged as an adjunct on the basis of syntactic iteration.

3.1.2.3 Head-dependence and head selectivity

A frequently applied syntactic generalization is that arguments are dependent on the verb (hence head-dependant) which results in their co-occurrence with a very restricted number of verbs, while adjuncts can appear with a significantly larger set of verbs.

This contrast is reflected in the examples in (8) below.

- (8) a. Mary {ate/smiled/broke a glass/saw the dog/stabbed the burglar} in the kitchen yesterday morning.
b. She {warned/*heard/*kicked/*feared/*smiled} her friend of the risk.

Very similar to the idea of *head-dependence* (Schütze 1995) presented above is the notion of *head selectivity* (Koenig et al. 2003), which presumes that verbs select the syntactic form of their complements, while the syntactic form of the modifiers is independent of the main verb in the clause. Thus, the verbs in the examples in (9) are assumed to select for a particular preposition to head the complement expressed by the prepositional phrase in the sentence.

- (9) a. You can always count *on him*.
b. She looked *at the picture*.

In contrast, in the examples in (10) it is suggested that the verbs simply allow for a prepositional phrase, marked with the preposition *with*, which encodes for an instrument.

- (10) a. Jim ate the dessert *with a spoon*.
b. Joe opened the tin can *with his knife*.
c. She cut the cake *with his knife*.

Although this criterion may help distinguishing whether a participant is an argument as in the examples in (9), it does not provide a basis for determining whether

the participant is in fact an adjunct. This is clearly demonstrated in the examples in (10) where a distinction should be made between verbs which may simply co-occur with an instrumental as in the examples in (10a) and (10b) and those which denote a situation that includes an instrument as the example in (10c).

In addition, we should take into account that in most of the cases these syntactic diagnostics are actually determined by the semantics of the verbs. Thus, for example, Koenig et al (2003) argue that the non-locational use of the preposition *on* as in the example in (9a) above occurs with a number of verbs which naturally fall into a class of their own. These are all verbs which denote a relation of dependence.⁷

Thus, it becomes apparent that the syntactic criteria proposed so far in the literature cannot serve as a reliable diagnostics for the argument/adjunct distinction.

3.1.3 Semantic Criteria

Koenig et al. (2002, 2003) suggest that since argumenthood cannot be equated directly to syntactic obligatoriness or overt syntactic realization we should look for semantic criteria instead. They propose two semantic criteria, namely *semantic obligatoriness* and *verb class specificity*, which jointly determine the argument status of a participant.

In addition, the authors argue that these criteria are based on semantic properties on which language learners rely in determining the participant information that is included in the representation of lexical items. Thus, both factors are directly observable and do not involve as a prerequisite the knowledge of the argument/adjunct distinction, which is used in many of the syntactic diagnostics suggested. Therefore, *semantic obligatoriness* and *verb class specificity* can jointly provide a basis for independent learning of the distinction between arguments and adjuncts.

⁷ The example given in Koenig et al. (2003) is in fact with the verb *rely*, which additionally proves that both verbs show similar syntactic behaviour based on the similarity in their semantics.

3.1.3.1 Semantic obligatoriness

The notion of *semantic obligatoriness* as defined by Koenig et al (2002, p.226) refers to the information that is “*entailed to hold of the class of situations denoted by a word.*”

Thus, they follow the suggestion of Dowty (1982) who proposed *entailment* as test for argumenthood as demonstrated in the examples in (11) below.

- (11) a. John sold the house *to Peter*.
b. Mary baked a cake *for her friends*.

While the phrase *to Peter* in (11a) is said to be entailed by the situation denoted by the verb *sell*, the event of baking does not entail the presence of a benefactor, such as the phrase *for her friends* in (11b).

Although entailment has been criticized (for example in Bresnan 1982) as an insufficient diagnostics for argumenthood, Koenig et al. (2003) believe that *semantic obligatoriness* represents merely one of the criteria needed for the proper account of the argument/adjunct distinction, namely, that it is indeed a necessary but not a sufficient condition. This can be easily demonstrated with the use of phrases denoting manner, time, or location, as all events possible in this world do entail in general a certain time when they have occurred (or will occur), as well as place and a definite manner, which may be specified or not in the utterance.

3.1.3.2 Verb class specificity

The second criterion, which together with semantic obligatoriness constitutes a necessary and sufficient diagnostics for the argument status of a participant, is *verb class specificity* (Koenig et al., 2002). This means that the information supplied by an argument must be relatively specific to the particular verb.

The joint application of the two criteria is thus illustrated in the examples in (12) below.

- (12) a. She **cut** the paper *with the scissors* yesterday evening.
 b. They **drank** their cocktail *with a straw* yesterday evening.
 c. He **put** the plates *on the table* yesterday evening.
 d. He **said** the prayer *at the table* yesterday evening.

While *cut* always describes a situation where an instrument is included, *drink* only allows an instrument to be included in some sub-types of the situations denoted by it. Thus, in the example in (12a) the instrument phrase *with the scissors* is both *obligatory and specific* for the verb *cut*, but one need not use an instrument to *drink*. Similarly, only the location phrase *on the table* is obligatory and specific for the verb *put* in the example in (12c), as opposed to the phrase *at the table* modifying the verb *say* in (12d). The time referring phrase *yesterday evening* is neither obligatory nor specific for any of the verbs in (12).

Therefore, an instrument should be included in the lexical representation of *cut*, and need not be present in the encoding of *drink*, as well as location should be encoded in *put* but not in *say*. *Time* referents, however, do not meet the requirements to enter the semantic representation of any of these lexical entries.

3.1.4 The Semantic Criteria at Test

To test the correctness of the criteria proposed, Koenig and his colleagues conducted several experimental studies, which I will briefly present in the next two sections.

3.1.4.1 Tests targeting the semantic obligatoriness condition

The notion of semantic obligatoriness discussed in Koenig et al. (2002) is further defined in Koenig et al. (2003) as the *Semantic Obligatoriness Condition* (SOC), which is repeated in (13) below.

- (13) *If x is the filler of an argument participant role r associated with predicate P, then any situation that P felicitously describes includes the referent of x.*

This condition was tested in a study aiming at determining whether semantically obligatory information associated with syntactically optional constituents may have an immediate effect on parsing of filler-gap sentences. The target verbs in the test semantically required an instrument (as the verb *stab* in the examples⁸ in (14a) and (14c)) while the control verbs (like *intimidate* in (14b) and (14d)) simply allowed an instrument in the particular event they described.

- (14) a. What type of weapon | did the knight | *stab* | the fiery dragon *with* [gap] in the famous story?
 b. What type of weapon | did the knight | *intimidate* | the fiery dragon *with* [gap] in the famous story?
 c. With what type of weapon | did the knight | *stab* | the fiery dragon [gap] in the famous story?
 d. With what type of weapon | did the knight | *intimidate* | the fiery dragon [gap] in the famous story?

Thus, in the examples in (14a) and (14b) the *wh*-NP fillers are syntactically possible as either a direct object or as indirect object, but semantically they can be realized only as the prepositional object of the stranded preposition *with*. Sentences like those in the examples in (14c) and (14d) were used as control sentences as the *wh*-PP fillers are straightforwardly syntactically impossible as a direct object and therefore must be associated with a later prepositional object gap.

The prediction was that the sentences with *wh*-NP fillers would elicit processing difficulties as the readers would first try to integrate the filler as the direct object of the developing sentence, unless the verb contains some participant information which leads to the assumption that the filler could be possibly associated with a later prepositional object gap. Thus, sentences with verbs which lexically encode an instrument are

⁸ The examples in (14) are re-created from the list of stimuli for experiment 2 in Koenig et al. (2003).

expected to display faster processing times than those that simply permit an instrumental phrase.

The results showed that the residual reading times to *wh*-NP filler sentences with *stab* verbs were similar to the *wh*-PP control sentences at any region. As expected, the *wh*-NP filler sentences with *intimidate* verbs induced longer residual reading times than their respective *wh*-PP control sentences in two of the initially defined regions, namely the verb and the direct object regions. These findings sustained the prediction that participant responses will differ for verbs that encode instruments and verbs that merely allow for an instrument to be present in the situation expressed by the sentence at hand, thus showing a difference in the participant information encoded in verbs like *stab* versus those like *intimidate* in the illustrative examples in (14).

Hence, the results from the *wh*-filler gap study confirmed the hypothesis that semantically obligatory information is encoded in the lexical entries of verbs and therefore plays an important role in the immediate representation readers develop for sentences thus influencing their on-line parsing.

In addition, these findings demonstrated that argument status cannot be assigned on the basis of participant category, such as instrument participants for example, as they cannot be uniformly defined as arguments or adjuncts. Instead, it is the semantics of verbs that determines whether a particular participant is part of the event denoted by the verb and therefore should be considered as an argument of this verb or not.

3.1.4.2 Tests targeting the semantic selectivity condition

The *Semantic Selectivity Condition* (SOC) (Koenig et al. 2003) builds on the notion of verb-class specificity and is presented in (15) below.

- (15) *A semantically obligatory participant role r for the situation denoted by verb V constitutes an argument if the role r is specific to V and a restricted class of verbs/events.*

This criterion was experimentally confirmed in a quantitative survey of the English verbal lexicon using the MRC psycholinguistic database (Coltheart, 1981). Following the SOC the authors predicted that a given participant role will be displayed with an apparent decrease in the percentage of verbs which lexically encode this participant information as compared with those that do not encode it.

Since the *agent* role is traditionally accepted as the most frequently occurring and unambiguous instance of argumenthood, Koenig et al. (2002, 2003) purposely selected to test agents in order to determine the limiting case for class selectivity. The participants in the tests were asked to assess whether the subject in the experimental sentences was involved in a situation that included cause or change of state, whether this participant was necessarily volitionally involved, and whether she/he had a mental representation of another participant in the situation. Thus, the subject was characterized by three independent agent properties – *causal force*, *volition* and *notion* (defined following Dowty 1991 and Davis and Koenig 2000), as present in the situation denoted by the test sentences. These were identified and rated separately following the concept of *semantic role* in Dowty (1991) which is defined as a cluster category and can be viewed as a complex of properties.

The results of the survey revealed that even the most frequent agent property *causal force* is restricted to slightly less than 30% of the English verbs. This finding supported the suggestion made by Koenig et al. (2003, 2003) that lexically encoded participant information is restricted to the verb and a very limited class of verbs/events.

In a subsequent survey targeting *instruments* and *participant locations*,⁹ the raters were asked to judge for each of the verbs whether it semantically required or simply allowed the presence of the relevant constituent (instrument, participant location, external location, or time) in all situations it felicitously described. The results showed a sharp drop in percentage of verbs judged to require the tested semantic roles with *times* and *external locations* being very frequent (99.8% and 98.2% respectively) and *instrumentals* and *participant locations* being very restricted (with the corresponding

⁹ The notion of participant location as used by Koenig et al. (2002, 2003) refers to location which indicates the place of the participant as result of the event denoted by the verb, as opposed to *event/external location* – the location at which the event takes place. In my research, the *participant location* is captured by the *End of Path* feature.

12% and 7%). Thus, these findings confirmed the hypothesis that obligatory instruments and participant locations are part of the lexical representation of a very restricted set of verb classes.

Taking this hypothesis as a starting point for their next study, Koenig et al. (2002, 2003) suggest that if participant information is lexically encoded in the semantic representation of a particular verb it is more likely to become activated upon recognition of that verb and be used in providing a sentence continuation. Thus, for a sentence continuation experiment they predicted that if the sentence contains a verb which requires an instrument, it will elicit more instrument continuations than if the main verb of the sentence simply allows for an instrument (as it was illustrated in the examples in (12a) and (12b)).

The results confirmed the prediction as they demonstrated that verbs which semantically require participant location or an instrument received significantly more completions of the respective type than expected by chance. There was also a significant difference between instrument/participant location completions and all other types of completions elicited by the targeted verbs. These results from the studies supported the results from the survey and together they can be used as reliable empirical evidence that participant information, which is lexically encoded, and is both semantically obligatory and relatively specific to the particular verb or a semantically restricted verb class.

3.2 Verb Sense and Subcategorization Preferences

Hare, McRae & Elman (2003) argue that verb semantics is one of the potential non-random factors underlying the verb's overall subcategorization preferences and specifically that some aspects of verb subcategorization information are not encoded in relation to the particular verb, but rather to the verb's specific senses. Thus, although verbs may occur with multiple subcategorization frames, the probability of these occurrences differ by verb sense.

In the following sections (3.2.1 and 3.2.2) I will briefly describe the notion of *verb sense* as defined by Hare et al. (2003) and the correlation between verb sense verbs subcategorization biases. Then in the next two sections (3.2.3 and 3.2.4) I will present

the results from the corpus analyses and the experiments carried out by the authors to test their hypothesis.

3.2.1 *Verb sense*

The notion of *verb sense* is used by Hare et al. (2003) to account for the different meanings of a polysemous word as opposed to the distinct meanings of homonyms like *mean* ('intend', 'unkind', 'average'). Thus, *verb sense* refers to highly related meanings, which are commonly distinguished in that usually one of them is the concrete physical sense and the other is an extended and more abstract sense of the verb (cf. Lakoff, 1987). This is illustrated in the examples in (16) below.

- (16) a. Kathrin felt *the rain on her face*.
b. Kathrin felt *that something was wrong*.

The sentence in (16a) employs the physical sense of the verb *feel* and displays one subcategorization frame [NP V NP], while in the sentence in (16b) *feel* is used with a more abstract or metaphorical meaning and displays a different subcategorization frame [NP V that_S]. Thus, Hare et al. (2003) argue that verb subcategorization profiles vary depending on the verb sense and that language users are sentient to that diversity and use this information in the parsing of temporally ambiguous sentences. Such ambiguities arise in sentences like the example in (17b) below.

- (17) a. Kathrin felt that *the rain was going to continue for days*.
b. Kathrin felt *the rain* *was going to continue for days*.

The post-verbal noun phrase *the rain* in (17b) is structurally ambiguous. It could be processed as either the direct object (DO) of *feel* (as it was in the example in (16a)) or the subject of a sentence complement (SC) as it is apparent in the example in (17a). This situation arises because the complementizer *that* is optional and may be omitted.

Therefore, it is not before the introduction of the verb in the SC (also called the disambiguation region) that this structural ambiguity can be solved.

3.2.2 *Verb subcategorization biases*

Although many verbs appear to display almost equal tendency in taking a DO or an SC, most of the verbs do show a bias when the probability of occurring with the one or the other syntactic structure is assessed in relation to the verb sense. Thus, Hare et al. (2003) suggest that verbs exhibit sense-contingent subcategorization biases (DO-biased or SC-biased verbs are distinguished in this particular study) and that this knowledge may play a role in the processing of temporally ambiguous sentences.

This proposal is also in line with earlier findings by MacDonald (1993) who has argued that the resolution of structural ambiguities is not only similar to lexical ambiguity resolutions but in many cases structural ambiguities are related to some sort of lexical ambiguity.

3.2.3 *Corpus Analyses*

For a set of candidate verbs Hare et al. (2003) carried out a multi-corpus-analysis investigating the correlation between overall and sense specific subcategorization preferences of a verb. Thus, twenty verbs, which could take both a DO and an SC and have been previously judged as clearly exhibiting different senses,¹⁰ were analysed in four different corpora: the Wall Street Journal (WSJ), Brown Corpus (BC), WSJ87/Brown Laboratory for Linguistic Information Processing (BLLIP), and Switchboard (SWBD), where the first three are written corpora and the last one is a conversational corpus.

The general subcategorization analysis showed the overall domination of the DO structure. However, the analyses of the subcategorization preferences contingent on the

¹⁰ All the verbs included in their research have been formerly categorized in WordNet (Miller, Beckwith, Fellbaum, and Miller 1990) as having several different senses.

verb sense demonstrated clear DO/SC biases thus pointing to a probabilistic correlation between verb sense and subcategorization profile.

3.2.4 *Empirical data*

To test whether language users are sensitive to the probabilistic relationship between verb meaning and syntactic behaviour, Hare et al. (2003) conducted four norming experiments and one self-paced study, which I present briefly in the following sections.

3.2.4.1 Off-line norming experiments

The first norming experiment was aimed at measuring the sense contingent subcategorization preferences out of context and in addition at establishing a baseline for the tendency of each verb to be used in either sense. The participants were asked to complete the target sentences which consisted of a pronoun and a verb in past tense only, as shown in the example in (18) below.

(18) He felt _____

Thus, the participants were not constrained in any way and no particular completions were elicited through surrounding context or by any other means.

The results of this study showed that out of context the verbs were non-significantly biased towards the DO sense. However, the overall tendency, disregarding the verb's sense, showed a significant preference of DO completions (62% vs. 20% with an SC structure).

The second experiment was designed to estimate the degree to which a specific context can promote a particular verb sense as well as to assess the sense-contingent verb subcategorization preferences. The target sentences were preceded by one-sentence contexts promoting either the DO-biased or the SC-biased sense of the verb as displayed

previously in the corpus analyses. An example¹¹ is given in (19) below, where (19a) presents a SC-biasing context and (19b) – a DO-biasing context.

- (19) a. (*SC*) Rick was snug inside the cabin, but his horses were outside for the night and that worried him. He felt ____
- b. (*DO*) Rick was beginning to get a little cold as he climbed the icy mountain. He felt ____

The findings revealed a significantly higher percentage of SC completions when the SC-biased sense of the verb was used, although DO completions were generally possible (71% of SC vs. 15% of DO structures). Similarly, there were more DO completions when the DO-biased sense of the verb was promoted, except that in this case the SC completions were extremely rare (89% vs. only 3%).

In addition, a comparison with the results from experiment 1 (where no context was included) showed that the SC contexts influenced more strongly the sense-contingent structural biases of the verbs than the DO contexts did.

The third experiment included the same target sentences, which contained an additional post-verbal NP, aiming at testing whether the NP would alter the perceived sense of the verb and thus its subcategorization preferences.

- (20) a. (*SC*) Rick was snug inside the cabin, but his horses were outside for the night and that worried him. He felt the weather ____
- b. (*DO*) Rick was beginning to get a little cold as he climbed the icy mountain. He felt the weather ____

Although the introduction of a possible NP did increase the DO completions in the DO-biased sentences, the decrease of SC continuations following SC-biased contexts turned out to be insignificant (from 89% to 84%). Thus, the including of an NP

¹¹ The examples in (19) through (22) are reconstructed from the list of stimuli given in Appendix B in Hare et al. (2003)

could not entirely hinder the role of the SC-biasing context, because it was also perceived as a plausible SC subject.

The fourth experiment was aimed at assessing whether the contexts alone were influential enough to promote the expected sense of the verb. Therefore, the sentences to be completed ended at the subject pronoun and no verb was included as shown in the examples in (21).

- (21) a. (*SC*) Rick was snug inside the cabin, but his horses were outside for the night and that worried him. He ____
b. (*DO*) Rick was beginning to get a little cold as he climbed the icy mountain. He ____

The results showed that although the contexts did not elicit many completions involving the specific target verbs, the continuations provided by the participants in the experiment generally described a scenario consistent with the context and the target sense of the verb, which proved that the contexts provided had in fact played their role.

The overall results from the off-line norming experiments confirmed the findings from the corpus analyses that there is a reliable correlation between verb sense and subcategorization probabilities and that context can bias language users towards a specific sense of a verb thus eliciting completions with particular syntactic structures that are contingent on the verb sense.

3.2.4.2 On-line reading experiment

To test whether language users actually employ the information of verbs' sense-contingent subcategorization probabilities (attested in the previous experiments) Hare et al. (2003) designed a moving-window self-paced reading experiment. The experiment was aimed at determining whether this knowledge would influence the resolution of temporary DO/SC ambiguities during on-line processing of sentences.

However, the sense-contingent subcategorization preferences of the verb are only one of a set of probabilistic constraints that may have effect in the test. In addition,

several other constraints are taken into account when analysing the results. One of them is the general tendency for the DO construction in English, attested in Experiment 1 described above. A second constraint regards the context influence on promoting a certain type of event, which is associated with a particular sense of the verb and therefore elicits specific syntactic structures, as seen in Experiments 2 and 3. A third one is dependent on the presence or absence of ambiguity (syntactically expressed by the respective absence or presence of the complementizer *that*) in the target sentences. Thus, in the unambiguous sentences the mere presence of the complementizer *that* is biasing towards an SC structure and the respective sense of the verb, while the ambiguous sentences favour the DO structure and trigger a possible DO sense interpretation of the verb.

Therefore, the model of the experiment predicts a combination of all available constraints (three of which were described above), which competitively influence alternative interpretations.

The participants were presented with a sentence long context (DO or SC biased) followed by the target sentence (ambiguous or unambiguous SC structure). Sentences had the same structure as in the examples¹² in (22) below.

(22) a. (SC) Rick was snug inside the cabin, but his horses were outside for the night and that worried him.

(Target) He felt₁ (that₂) the₃ weather₄ might₅ become₆ a₇ problem₈ | as the night wore on.

b. (DO) Rick was beginning to get a little cold as he climbed the icy mountain.

(Target) He felt₁ (that₂) the₃ weather₄ might₅ become₆ a₇ problem₈ | as time wore on.

¹² The example in (22) is taken from the Appendix B in Hare et al (2003). The indexes given to the words in the target sentences are mine and are meant to indicate the different regions that were measured by the authors of the study.

The subjects had first to read the entire context sentence and press a response button thus indicating they have finished reading. Then the target sentence was revealed one word at a time, as the subjects had to press the button in order for the next word to appear, while the previous one disappeared. Thus, the reading times for the exact region of each target word were measured.

The results are described by Hare et al. (2003) following a constraint-based satisfaction account developed in earlier works by Spivey & Tanenhaus (1998) and McRae et al. 1998). First, the authors predicted which constraints should be active at each word region and how they could affect on-line processing. Thus, they suggested that if the constraints available to the readers sustain a single interpretation, little or no ambiguity effects should be detected. On the contrary, if there are several competing constraints supporting alternative interpretations, this should lead to larger ambiguity effects. These predictions were fully confirmed by the reading times measured for the different regions during the on-line experiment.

Furthermore, an unexpected context effect in the verb region was observed. Reading times for the main verb (*felt* in the example in (22)) were significantly shorter following DO contexts. Among the indications pointing towards this possibility the authors include findings from previous experiments, which showed that target verbs were in general DO-biased in meaning (although it was not significant) and that verbs were more easily interpreted following DO-contexts. These differences may also be based in the correlation between the concrete, physical sense of the verb and its subcategorization preferences, as it was usually the DO sense of the verb that was more concrete.

The overall results thus showed that participants were strongly influenced by the context which promoted a particular sense of the verb thus leading to sense-contingent structural expectations. Consequently, reading times were significantly longer following DO-biasing contexts (as in the examples in (22b) above) versus those following SC-biasing contexts (the examples in (22a)), as the DO-biased contexts mislead the subjects to expect a DO structure with the verb to follow, while all the target sentences in this study used the verbs with a SC-structure.

In addition, Hare et al. (2003) suggest how data from the experiments may be incorporated in the theory of lexical representations. On their account the results are

most consistent with the view of distributed attractor networks,¹³ which treat each sense as a separate but a nearby basin of attraction. A number of factors play a role in determining the distance between those basins in the dimension of the semantic/syntactic space. These include semantic overlap, structural considerations (presented as trajectories into and out of the attractor basins), as well as phonological and orthographical similarities. Thus, the information of sense-contingent subcategorization preferences of a verb is presented as the probability of moving along paths formed by the various trajectories coming out of the attractor basin representing the particular verb sense. Then the role of the context in the experiments is seen as manipulating the interpretation to start from the intended attractor basin of the verb.

Finally, based on the results from all the experiments, Hare et al. (2003) conclude that verb meaning is one of the relevant factors underlying the on-line processing of temporarily ambiguous structures, as language users are sensitive to the specific sense of the verb and have structural expectations that are contingent on it. Thus, they interpret structural ambiguities in a manner coherent with the verb sense.

3.3 *The Sign Model and the Lexical Encoding of Verbs*

The formal analysis of the empirical data used in my research is based largely on a framework developed in works by Dimitrova-Vulchanova (1996/99) and Hellan & Dimitrova-Vulchanova (2000). On this account, lexical items are presented as *signs*, wherefrom the name of the framework *The Sign Model*. The notion of *sign* here is used very much in the Saussurean tradition employing the two aspects of the linguistic sign – the formal and the conceptual one.

In this section, I will present the *Sign Model* framework with respect to the lexical encoding of the verbs of interest. First, I will briefly describe the structure of the sign as introduced by Dimitrova-Vulchanova (1996/99). Then, in section 3.3.2, I will further specify the constitution of the *cell* which includes the semantic components of

¹³Presented in works by Elman (1991, 1995), Tabor, Juliano, & Tanenhaus (1997), Kawamoto (1993), and Klein & Murphy (2001) among others.

the sign. In section 3.3.3, particular attention is paid to the dimensional part of the *cell* which accounts for the various aspects from which a situation can be analysed. Section 3.3.4 illustrates the notion of Criteriality and discusses some of the elements which are defined as criterial in Hellan & Dimitrova-Vulchanova (2000). Finally, section 3.3.5 will address issues of the syntax-semantics interface as the relations between semantic participants and their overt morpho-syntactic realization is presented.

3.3.1 *The structure of the sign*

The structure of the sign revolves around two main dichotomies. The first one sets apart Meaning and Expression, thus distinguishing the semantic properties of a lexical item from its phonological and morpho-syntactic features.

The other dichotomy, *Figure* vs. *Gestalt*, reflects the distinction between the overt expression of a lexical item and the set of properties ascribed to the participants in the situation denoted by it. Thus, *Figure* represents some formal basic properties of the item, such as word-class category and phonological form.

The *Gestalt* of a sign is further divided into two parts – a semantic part and a morpho-syntactic part. The morpho-syntactic part is labelled the *Frame* and describes the lexical item on two distinct levels related to grammatical function-structure (GF) and morpho-syntactic realization (MSR), respectively.

The semantic part also consists of two components, *Aspectual specification* and *Element/Dimension specification*, described in detail in section 3.3.2.

Thus, the structure of the sign may be generally illustrated as in Fig.1 below.¹⁴

¹⁴ Although slightly changed in appearance, the structure of the sign in Fig. 1 is essentially the same as the one given in Dimitrova-Vulchanova (1996/99).

Gestalt	Conceptual/Semantic Representation (Cell)		Morpho-Syntactic Structure (Frame)		
			GF-structure	MSR	
	Aspect	Global Specification		morpho-syntactic devices	
		primitive aspectual features			
Element Specification/ Dimensions	element 1 — element 2 — ... element n —	— GF 1 - — GF 2 - — GF n -	— X — Y — Z		
Figure	Sound Representation		Morpho-syntactic category		
	Phonetic Structure	Phonological Structure			
	[...]	/.../	X ⁰		

Fig.1 Structure of the *Sign*

The cell together with the phonological structure of a word are considered as the minimal sign of this word, while the expanded sign will also include the morpho-syntactic characteristics ascribed to the lexical item at hand.

In the subsequent discussions of the investigated verbs (chapters 4 and 5) I will use primarily the *cell* to represent the information that is lexically encoded in the verbs at hand. I further describe its structure and elements included in the representation in the following two sections.

3.3.2 The Cell

As already mentioned above, the semantic part of the *sign* is called *Cell* and consists of two parts defining the event on two separate levels¹⁵ – the *Aspectual specification* and

¹⁵ Similar approaches to the representation of verbs include the tiered representation (Jackendoff 1987, 1990, Nikanne 1990, 1995) presented in section 2.4 above and Grimshaw's Argument Structure (1990), to mention some.

the *Element/Dimension specification*. This is essential because verbs represent relations and therefore cannot be conceptualized alone, apart from the entities that participate in the situation described.¹⁶ Thus, different aspects of the event denoted by the verb need to be specified, as they are relevant for the truthful representation of the verb meaning and have reflexes on the mapping of conceptual items.

3.3.2.1 Aspectual specification

The aspectual component of the cell contains primitive features which describe the situation on an aspectual level, thus regarded as very similar to the notion of *Aktionsart*. This component specifies for a number of factors from the semantic representation of verbs given in (23) below.

- (23) a. *Situational* vs. *Non-Situational* – reflecting whether what is expressed by verb is situated in time or not (for ex. *receive* vs. *require*).
- b. *Dynamic* vs. *Stative* – relevant only for situational verbs and reflecting whether some kind of change or Force emission is involved or not (for ex. *hit* vs. *hate*).
- c. *Monodevelopmental* vs. *Non-Monodevelopmental* – depending on whether the dimensional part includes Monodevelopment (a monotonic process) or not (for ex. *push* describes a situation which includes a steady monotonic process resulting in a unidirectional change).
- d. *Protracted* vs. *Non-Protracted* – a contrast which is close to the traditional distinction ‘durational’ vs. ‘non-durational’ (for ex. *drag* vs. *stab*).
- e. *Definedness for End Point (DEP)* vs. *Non-definedness for End Point* – refers to the presence or the absence of a point/an entity with respect to which the situation can be measured as fulfilled.

¹⁶ For similar accounts cf. Langacker 1987 and Dabrowska 2006, among others.

f. Completed vs. Non-Completed – signifies whether the situation in its entirety is perceived as completed or not. DEP is defined as a sufficient, but not necessary condition for Completedness to obtain. Thus, the set of situations marked for DEP can be viewed as a subset of the set of situations marked as Completed.

The first four features described in a) through d) can be retrieved from the dimensional specification of the verb, as they relate to the actual conceptualization of the event denoted by the verb. The last two features, (e) and (f), are characterized as Global and are defined on the interface between semantic properties of the verb and its morpho-syntactic realization.

3.3.2.2 Element Specification

The *Element specification* part consists of a number of *dimensions*, each of which represents a different aspect of the event encoded by the verb. Thus, every participant (element) is assigned a set of features that reflect the various sides of its involvement in the event or any sub-event that is part of the main event.

With respect to their importance for the lexical representation of verbs a more detailed picture of the different dimensions is offered in section 3.3.3 below.

3.3.3 The Dimensions

The function of the dimensions¹⁷ is to present distinct but concomitant aspects of the situation denoted by the verb. Here the main idea is to recognize the multi-dimensional specification of participants in the situation of hand, as well as the differentiation of the sub-events constituting the main event. Each of those sub-events should be described separately in detail. All of them describe the situation as a whole.

¹⁷ The concept of the dimensions is similar to the view of tiered representation (cf. Jackendoff 1990, Nikanne 1990, also discussed in Chapter 2 in this work).

A similar formal proposal on event structure representation that includes segmentation into separate events was introduced by Pustejovsky (1995) and briefly presented in section 2.3.1.2 above. Thus, various relevant aspects of the situation are accounted for with respect to their significance for the mapping of participants from the semantic representation of the verbs to their syntactic realization.

Each of the dimensions may consist of one or more values. An element does not need to be assigned a value in each dimension, nor should a dimension be present in the *cell* if it has no values, i.e. if it is not present in the situation denoted by the particular verb.

Four of the most relevant dimensions, with respect to the set of verbs at hand, are presented in sections 3.3.3.1 to 3.3.3.4 below.

3.3.3.1 Force

One of the most relevant dimensions for the verbs examined in this project is the dimension of *Force*. It represents a situation characterized by an emission of physical force. Thus, it may incorporate the values of *Source* (the participant releasing the force), *Source Extension* (the part of the participant, if any, performing the action), and *Absorber* or *Limit* (the item upon which the force has been performed).

The distinction made between *Absorber* and *Limit* reflects the possibility of the *Absorber* to undergo a process or a movement as result of the force applied to it; while *Limit* is regarded as the last entity in the *Force* chain, and no such movement is to occur. Thus, an element with the value of *Absorber* may also be co-indexed with the value of *Monodeveloper* and hence be a part of a *Monodevelopment*.

Thus, in both sentences in (24) *Peter* bears the value of *Source*, and the value of *Source extension* is assigned to *his bare fist* in (24a) as well as in (24b). However, while *the wall* in (24a) receives the value of *Limit*, in the sentence in (24b) *the mirror* carries the value *Absorber*.

- (24) a. Peter hit the wall with his bare fist.
 b. Peter broke the mirror with his bare fist.

Moreover, each of these values does not independently define the participant to which it is assigned. It is usually only one of the set of values characterizing co-indexed elements in the various dimensions. Thus, each participant is defined by the set of values ascribed to it within the different dimensions and participant identity is ensured through co-indexation. Therefore, the participant expressed by *mirror* in the example in (24b) above should be characterized so far by a set of co-indexed values [Absorber_n, Monodeveloper_n], where each of the values is ascribed to the participant in a different dimension, and **n** is the index which is used to denote this particular participant throughout the various dimensions.

3.3.3.2 Monodevelopment

The dimension of *Monodevelopment* (short for ‘monotonic development’) is present when at least one of the participants in the event is involved in a monotone¹⁸ unidirectional process. That is, if a participant undergoes a process of progressive changing with respect to a certain parameter, the value of *Monodeveloper* can be ascribed to it. The process is regarded as monotonic when it can be presented as consisting of a sequence of stages with successive values representing the progressive change of the parameter at hand. This parameter could specify a change in the integrity of the participant as for example with *break* in the sentence in (25a), or it could involve a change in its quality like exemplified by *melt* in the sentence in (25b). The most common case is change of location in space (either along the vertical or along the horizontal axis) as with *fall* and *drag* in the examples in (25c) and (25d) respectively.

- (25) a. Jonathan broke the window.
 b. The ice-cream melt.
 c. The child fell on the grass.
 d. Susan dragged her suitcase into the room.

¹⁸ Used in the mathematical sense of the word.

Thus, the dimension of Monodevelopment includes the value of *Monodeveloper* (the entity undergoing the monotonic development) and *Medium* (the parameter with respect to which the Monodevelopment is specified), such as *Integrity*, *Location* (mainly regarding *path*), and *Quality* as exemplified in the examples in (25) above. Hence in the example in (25a) *window* is *Monodeveloper* with *Medium: Integrity*; *ice-cream* in the example in (25b) has a *Medium: Quality*; and *child* in (25c) and *suitcase* in (25d) are both *Monodeveloper* with *Medium: Location*.

Then following the discussion in the previous section 3.3.2.2, we can describe for example *suitcase* in (25d) as the entity **n** defined by the set of values [Absorber_n, Monodeveloper_n].

3.3.3.3 Conditioning

For many verbs, a further dimension of *Conditioning* is possible in close relation with *Monodevelopment*. Conditioning applies when, in a given context, a given event or actor, called the *Conditioner*, is sufficient to release a certain event, called *Conditioned*.

This event is usually a *Monodevelopment*. For example, in the sentence in (26a) below *Kate* is the *Conditioner* for the event of breaking the glass.

- (26) a. Kate broke the glass.
 b. The glass broke.

In contrast, in the sentence in (26b) no Conditioner is identified and no Conditioning obtains in this usage of *break*. Thus, *glass* in (26a) is defined by the set of values [Conditioner_n, Monodeveloper_n], while in the situation exemplified in (26b) only the value of *Monodeveloper* can be ascribed to *glass*.

3.3.3.4 Control

The notion of *control* is applicable to situations where one participant is in command of the situation (or a conditioned event) in the sense that it can terminate the situation or guide its course. Thus, a participant that combines the roles of *Conditioner* and *Controller* is referred to as *Initiator*.

As already mentioned in section 2.3.1.2 a lexicalized event can have at most a binary structure, which can be simply presented as consisting of two sub-events. Thus, although in nature sequences of several events in a chain are possible, the situations that are lexicalized by verbs can include at most two sub-events.

This is also true for African and Asian languages some of which have the syntactic phenomenon dubbed *serial verb construction*, where an event is described using several verbs organized in a chain within a single clause. The verbs share a subject (usually the actor) but they may also have arguments on their own, as illustrated in the example in (27) below.¹⁹

- (27) Aba yé-è asɔr má-à Kofi (Fante)
 Aba do-COMPL prayer give-COMPL Kofi
 Aba prayed for Kofi.

Thus, each of the verbs represents a single sub-event and again no single verb can lexicalize more than two sub-events. (cf. Osam (2003) and Kropp Dakubu (2003) for more information on the nature of the serial verb constructions, and Dimitrova-Vulchanova & Martinez (forthcoming) on cases of motion encoding in verbs in Akan.)

With respect to the components of the force dimension the two lexicalized sub-events can be viewed as **movement** and **contact**. Thus, Dimitrova-Vulchanova (1996/99) suggests that there are only two possible sequences: A) movement – contact, or B) contact – movement.

¹⁹ The example in (27) is adopted from Osam (2003). It is in Fante, a dialect of the Akan language, spoken in Ghana.

In addition, the notions of *minimum* and *maximum* control are introduced. Thus, an instance of minimum control in a situation of type A) is presented with '*kick the ball*,' while minimum control in type B) situations is illustrated by '*throw the ball*.' What these situations have in common is that the scope of control pertains to the first sub-event in the event chain and is exhausted with the release of the second sub-event, without any control over it.

Maximum control is demonstrated with the situation in '*shoot the president*' thus showing that Initiator in a situation of type B) has control over the result of the movement in the second sub-event.

Hence, we can generally conclude that the scope of control may pertain to the first sub-event only, in which case it is defined as a minimum control, or extend to the second sub-event. The latter is the case of maximum control.

In line with this proposal, I will suggest that there are four possible situations which may arise from the combination of the two sequences given in Dimitrova-Vulchanova (1996/99) and the application of the notions of minimum and maximum control,²⁰ The situations enlisted in (28) below are illustrated respectively by the four sentences in the following examples in (29).

- (28) a. movement – contact, minimum control
 b. movement – contact, maximum control
 c. contact – movement, minimum control
 d. contact – movement, maximum control

- (29) a. He kicked the ball.
 b. He kicked the ball into the goal (to score a point).
 c. She threw his clothes.
 d. She threw his clothes into the laundry basket (to wash them later).

²⁰ Although it is argued in Dimitrova-Vulchanova (1996/99) that the scope of Control is defined by the entity expressed as the direct object and prepositional phrases like '*to John*' in '*throw the ball to John*' stay outside the scope of control, I will suggest that the four situations proposed in (28) can in fact be recognized in the overt morpho-syntactic realization of the verbs as shown in the examples in (29).

The grammaticality of the intentional phrases in parentheses suggest that the participant expressed by the subject in the examples in (29b) and (29d) has control over the situation and its outcome. Thus, although the verbs lexical representations do not specify for any result of the second sub-event, the interaction between verb semantics and sentential semantics implies an extended scope of control. This phenomenon can be viewed as a language device to allow for verbs to denote some extra-linguistic situations which implicitly include a third sub-event by expressing its result. Thus, the examples in (29b) and (29d) may be roughly paraphrased²¹ as the sentences in (30a) and (30b) respectively.

- (30) a. He kicked the ball and it went into the goal.
 b. She threw his clothes and they went into the laundry basket.

Apparently, a third sub-event takes place in the situation and it cannot be part of the information encoded in the main verb. Instead, it is only implicitly present, marked by the *End of Path* phrase PP_{into}.

Thus, I will suggest that the maximum scope of control in situations of this type is realized on the level of Global aspectual specification, if the situation is marked by DEP, that is, defined for an end point. Hence only *End of path* phrase may extend the scope of control, but not *Path orientation* phrase (like PP_{to} for example) as seen by the ambiguous interpretation of the result in a situation like in '*threw the ball to John.*'

This proposal is also in line with recent analysis on the Directional Modality Construction in Finland-Swedish (Nikanne & Östman 2006) where Nikanne explains the unexpressed directional verb (like *åka* 'drive', *gå* 'walk', *resa* 'travel') in sentences like *Marit måste til Åbo* ('Marit must go to Turku') with an implicit relation between the Theme and the end point in order to connect the theme to the rest of the situation.

Thus, I would suggest that cases like the ones presented in (29b) and (29d) may be analysed as a causative variant of the Directional Modality construction, whereby the

²¹ However, the sentences in (30) are not conceptually equivalent to their counterparts in (29) as in (30) the event denoted by the verb in parentheses is somewhat disconnected from the main event and happens 'incidentally,' which is not the case in the examples in (29) above.

overt realization of an *End of Path* phrase extends the scope of control of the main verb although the implicit relation between the Absorber and the End point is not part of the lexical representation of the verb at hand. Therefore, this additional sub-event must not be included as part of the cell of the lexical item, but may be present in the expanded sign of the verb as used to denote the situation at hand.

3.3.4 Criteriality

As already suggested above, every participant in a situation is defined by the set of values characterising **co-indexed elements** in the different dimensions. Furthermore, the meaning of a verb is identified with the conditions that have to be met by the participants in a situation so that it can count as being expressed by this particular verb.

The notion of *criteriality* then, presented in Hellan & Dimitrova-Vulchanova (2000), applies to the items of a cell that have properties by which the situation is **uniquely** identified as belonging to a certain type. Thus, *criteriality* is seen (ibid.) as one of the members of a set of lexical semantic factors which serve as the source for predictions of the morpho-syntactic environment of a verb based on its meaning.

The items defined as being *criterial* by Hellan & Dimitrova-Vulchanova (2000) are listed in (31) below.

- (31) a. An item with the value ‘*Monodeveloper.*’
 b. A *Source* whose Launch-part²²
 i) behaves monotonically, or
 ii) is specified for inherent properties.
 c. A *Limit* with sustained contact.
 d. An item characterized for *Posture.*
 e. A *Source* for an Iterative activity with a cumulative *Target.*

All these items specify properties which can uniquely identify a situation as belonging to a certain type and in that sense they are criterial for this situation.

²² I refer to this element as the *Source Extension*.

This proposal is in line with the suggestion made by Koenig et al (2002, 2003) as it does not consider the syntactic realization of a participant as the only possible criteria of the presence of an element in the conceptual structure and hence in the lexical representation of a verb.

On the opposite, the elements which are criterial may not be present as they are conceptually specific for the type of situation denoted by the verb at hand and need not be overtly present. Consider the examples in (32).

- (32) a. I ate my breakfast.
 b. I ate.
 c. I cut the cake.

Irrespectively of the syntactic realization of a direct object as can be seen in the examples in (32a) and (32b), there is no situation of this world that could be denoted by the verb *eat* and would not include an entity eaten. Likewise, the situations of this world that can be described with *cut*²³ do include the presence of a cutting instrument although it may not be syntactically expressed as in the example in (32c).

In fact, the elements, which are tightly incorporated in the meaning of the verb, often cannot be syntactically expressed, unless they convey additional meaning. This was first discussed in detail by Jackendoff (Jackendoff 1990, among others) and can be illustrated for English and Bulgarian with the examples in (33) and (34) below.

- (33) a. I buttered my toast *with butter/with French butter.
 a' Namazax si filijata s maslo/s frensko maslo.
 spread refl.toast with butter/with French butter
 b. Omaslih si rucete *s maslo/s maslinno maslo.
 buttered refl.hands-the with butter/with machine butter
 b' I oiled my hands *with oil/with lubricating oil.

²³ As discussed later in section 4.1, I distinguish whether a verb denotes a situation of concrete physical impact or an extended meaning is imposed. Thus, if the verb *cut* is used with an extended meaning, the situation may not include a tangible instrument as in "*cut a scene*." However, a means by which the cutting has occurred is still implicitly present.

Thus, the PP phrase *with butter* in (33a) cannot be overtly expressed unless it provides information that is not already incorporated in the verb *butter*, as in this case the adjective *French* in the phrase *with French butter*. The same noun can be expressed in the Bulgarian translation of (33a) as shown in the example in (33a'), as it does not use the verb *butter* but a more general verb with a meaning closer to *spread*. However, the Bulgarian example in (33b) which uses the verb *omaslja* (*butter*) demonstrates similar syntactic behaviour as its English counterpart.

Similarly, in the examples in (34) below English and Bulgarian verbs used in the same context differ in their syntactic patterns due to differences in conceptualization.

- (34) a. I watered my plants *with water/with rain/tap water.
 a' Poljax si rastenijata s voda/s dūždovna voda.
 poured refl.flowers-the with water/with rain water
- b. I flooded my bathroom with water/with the water from the laundry machine.
 b' Navodnix si banjata *s voda/s vodata ot peralnjata.
 watered refl. bathroom-the with water/with water-the from laundry-machine-the

The pairs of sentences in (34a) and (34a') and (34b) and (34b') respectively are translational equivalents but do not use corresponding verbs and therefore allow for variation in the overt realization of semantic participants. The English verb *water*, although different in meaning from the Bulgarian *navodnja* (*flood*), encodes the same semantic participant (*water*) and hence the similarity in the syntactic behaviour of the two verbs as seen in the examples in (34a) and (34b').

The process of syntactically unexpressed or also called implicit participants has been discussed by Jackendoff (1990) as a kind of direct object deletion on the basis of a *non redundancy condition* operating in cases of context specificity such as the incorporated Theme in denominal verbs.

However, the examples in (33) and (34) demonstrate that this phenomenon is not based on context only, but is related to conceptual specific information²⁴ encoded in the verbs.

3.3.5 *Realization of Criterial Elements*

In this section I discuss further the possible syntactic realization of elements in relation to their semantics. As it was briefly suggested in the previous section 3.3.4, criteriality plays an important role in the mapping of lexically encoded participants onto syntax. Thus, a rule is proposed in Dimitrova-Vulchanova (1996/99) which regulates the number of the criterial participants expressed overtly. It is called *Requirement for realization of criterial elements* and is replicated in (35) below.

- (35) If a verb has criterial participants, at least one of them must be expressed in a canonically retrievable position, or its conceptual presence must be confirmed.

The purpose of this rule is to specify the minimum number of elements that may be involved so that the verb can count as denoting a certain type of situation. In addition, the rule in (35) envisages predictions directly related to grammaticality.

In order to discuss the application of this rule I must first present the notions of canonical positions and canonically retrievable positions as introduced by Dimitrova-Vulchanova (1996/99).

3.3.5.1 **Canonical positions. Canonically retrievable positions**

The standard mapping of elements from the verb's semantic representation to the overt grammatical realization is called *canonical*. Thus, some of the most regular canonical

²⁴ cf. Dimitrova-Vulchanova & Welsgerber (2006) for a notion of context incorporating exactly this kind of information.

expansions align *Initiator*₁ [Source₁, Conditioner₁] with Subject and *Absorber*₂ with Direct Object as exemplified by the corresponding indices²⁵ in (36).

(36) [Michael]₁ hauled [his friend]₂ into the pub.

The Monodeveloper₁—Subject alignment is also considered as a canonical expansion in cases where *Monodevelopment* is the only dimension in the cell as in the example in (37) below.

(37) ... [they]₁ slapped against the wall. (BNC: BNU 1230)

A certain ranking hierarchy is observed in the sense that if a participant combines roles from several aspects of the situation denoted by the verb, it is regarded as a stronger applicant for the grammatical function at hand. This idea is very much in the sense of Dowty's proto-properties (1990). The grammaticality of the various possible expansions, however, is controlled by an axiomatic principle that ensures the **unique** assignment of the functions of Subject and Object (both direct and indirect object).

Non-canonical expansions are considered those which map for example *Absorber*₁ onto the function of Subject and *Initiator*₂ onto a prepositional phrase as illustrated in the sentences in (38) below.

(38) a. [He]₁ was stabbed [by a burglar]₂.
b. [The bread]₁ cuts well.

In close relation to the instances of non-canonical expansions is the notion of canonical retrievability as defined in Dimitrova-Vulchanova (1996/99). *Canonical retrievability* applies in cases where the original canonical expansion can be recovered directly with the help of different syntactic or semantic mechanisms.

²⁵ The indices in this and the following examples are used for clarity and do not implicate an order of importance amongst the participants.

One of the central grammatical mechanisms is called *morphological flagging* and is observed in sentences like the one in (38a) using passive morphology. Thus, the canonical positions of the Absorber (*he*) and the Initiator (*a burglar*) in (38a) can be retrieved by means of the passive grammatical marker which suggests that these are non-canonical positions. Hence, both realizations can count as canonically retrievable.

The semantic mechanisms involved in the process of canonical retrievability are incorporated in the meaning of the verb and rely heavily on the notion of criteriality. These are the processes of element elimination and element implication which are further described in next section under the common term suppression.

3.3.5.2 Suppression of elements and the survival of the criterial

When a participant is referred to as *eliminated* it means that it is neither syntactically realized, nor is perceived as implicitly present. This is the case with the *Initiator* in the intransitive usage of *break* as in *the window broke* or in the example in (26b) above. A necessary condition for a participant to be eliminated is that it is **not** criterial.

However, not every element that is present semantically in the situation denoted by the verb at hand has to be expressed syntactically. An element which is conceptually realized as taking part in the event but is not overtly realized in syntax is called *Implicated*. One such case of *Element Implication* is for example the participant marked as *Initiator* in a passive construction as illustrated in (38a) if the *by*-phrase is omitted. Another example of an implicated participant with the value of *Initiator* is observed in the cases of middle constructions as the example in (38b) above.

Implication of participant with the value of *Source Extension* can be found with verbs like *throw*, *cast*, *slap*, and *kick*, where the extension (*hand* or *foot*) in general is not syntactically realized.

A necessary condition for a participant to become implicated is that it is either criterial, or morphologically flagged. Thus, the generalization of the implicit presence of an element is described as the *Survival of the Criterial* axiom (Dimitrova-Vulchanova 1996/99) stating that *if a participant is criterial, it cannot be eliminated*.

3.3.5.3 Conceptually present participants

The conceptual presence of a participant is established based on fulfilment of one of the factors suggested by Dimitrova-Vulchanova (1996/99) and enlisted in (39) below.

- (39) A participant is conceptually present in a sentence if
- i. it is overtly expressed
 - ii. it is canonically retrievable
 - iii. it has a criteriality value attached to it

These criteria constitute an important part of the *Requirement for realization of criterial elements* given in (35) above, since their execution is particularly relevant for the interpretation of the suppressed *Initiator* in passives and middles (Dimitrova-Vulchanova 1996/99).

As it becomes evident from the discussion of the empirical data presented in chapters 4 and 5, the requirement presented in (35) above, together with the criteria defining the conceptual presence of a participant (stated in (36)), play an important role in the lexical representation of verbs including instrument or path referents. This is also in line with the semantic criteria for lexical encoding proposed by Koenig et al. (2002, 2003) and the results from their studies presented in section 3.1 above.

3.4 Conclusions

In this chapter, I presented three frameworks which have served as the theoretical background to the present work. These were used as helpful milestones in my research as they all explore the information that could be encoded in the lexical representation of verbs and its mapping onto syntax.

Section 3.1 described a relevant theoretical approach by Koenig, Mauner, & Bienvenue (2002, 2003) on lexically encoded participant information, together with the experimental studies conducted to test this hypothesis. The semantic criteria proposed and the methodology of the tests discussed in this section were used in the design and

the implementation of my own experiments aimed at mapping of semantic properties of verbs with displayed patterns of syntactic behaviour.

In section 3.2, I outlined a proposal by Hare, McRae & Elman (2003) stating that verb subcategorization preferences are contingent on verb sense. This was followed by brief a presentation of the results of their experiments carried out to test the correlation between verb sense and the verb's preferred subcategorization frames.

Section 3.3 introduced the main ideas of a formal framework currently known as *The Sign Model* (Dimitrova-Vulchanova 1996/99, Hellan & Dimitrova-Vulchanova 2000) which I have used as a theoretical model for the formal analyses of the empirical data in my research.

4. Corpus Data Analyses

As already discussed in the previous two chapters, this research follows the hypothesis that the meaning of a verb is revealed to the greatest extent in the way it is used by native speakers in language production and comprehension. Therefore, I aimed at conducting a comprehensive investigation of existing corpora resources, which can bring forth reliable evidence for an in-depth picture of verb semantics. With this in mind, I have analysed a sample set of verbs and examined their semantic properties with respect to their syntactic distribution and types of syntactic patterns displayed in the available corpora resources. This is in line with recent studies conducted by Koenig et al. (2002) and Hare et al. (2003) presented earlier in sections 3.1 and 3.2, respectively.

The corpus data used in this research have been collected from several main corpora. The Brown tagged corpus and the LOB tagged corpus have been used in the preliminary investigation for the English verbs. The results presented in this work are based on later searches in the British National Corpus (BNC) for English, and the Large Written Corpus of Bulgarian (hence LWCB) for Bulgarian.

The BNC has been preferred as it is much larger than the other two corpora mentioned above for English. The BNC is a collection of written and spoken language material from a wide variety of sources with the total amount of 100 million words. To collect my data, I have used the BNC online search service available at <http://www.natcorp.ox.ac.uk/>. As the corpus is tagged, I could easily restrict my search by part of speech. The Simple Search option showed the total frequency in the corpus and displayed up to 50 randomly selected results (different set of results was generated

each time a search was conducted) from all the occurrences found in the corpus, with a maximum of one sentence of context for each hit.

Being the only corpus of Bulgarian of an appropriate size, the LWCB has been used both for the preliminary investigation and as a basis for the final results of the Bulgarian data. LWCB is constructed at the Department of Computational Linguistics, The Bulgarian Academy of Sciences (<http://dcl.bas.bg>). The corpus comprises original and translated texts from diverse thematic domains and a variety of genres like fiction, journalese, government, and sports documents. It is the largest corpus of contemporary Bulgarian, with more than 35 million words.

This chapter presents the processes of collecting and analysing the corpus data, as well as an introduction to the terminology used (section 4.1) together with the results of the analyses reflected in the grouping of the verbs examined (section 4.2). Section 4.3 focuses on some of the extended uses of the verbs at hand as encountered in the corpus data.

4.1 Methodology of the analyses and terminology used

Aiming at a more detailed analysis of a specific set of verbs, instead of a shallow coverage of a wider range of verbs, I have selected several basic verb types in English and Bulgarian. Special attention was paid to approximately twenty verbs (listed in Appendix A), subgroups of what are called *Verbs of Contact by Impact* (as defined in Levin, 1993) along with verbs that include motion (in Levin's classification, those fall in the group of *Throw Verbs*). These were analysed for the types of semantic participants they display with respect to the syntactic patterns in which they occur.

In the following two sections, 4.1.1 and 4.1.2, I describe briefly the processes of collecting and analysing the corpus data used in this research. The results of the analyses are organized in several tables, separately for English and Bulgarian, given in Appendices B and C, respectively.

4.1.1 Methodology

For the purposes of this research, I restricted the searches to a set of approximately twenty verbs from English and their correlates in Bulgarian. These verbs were chosen as representatives of several semantically related verb classes that show a diversity of patterns of alternations allowed and a variety of constraints, which they pose on their syntactic environments.

For each of the analysed verbs whose frequency exceeded one hundred occurrences, I have created a set of randomly selected samples of one hundred sentences among all the findings in the corpus. This resulted in sets that were representative for the corpus and appropriate for examination in the research.

Because of their relatively low frequency, three of the Bulgarian verbs were represented with a smaller number of occurrences. These were *capna* (*smash*, where only contact is implied) – 44, *rŭgna* (*stab*) – 44, and *draskam* (*scratch*) – 50.

The search included all the verb forms, which were specified as a list of words for Bulgarian, and searched by tag identification for English. The English tag-set used in the research consisted of the following tags: VVB (the finite base form of lexical verbs), VVD (the past tense form of the lexical verb), VVG (the *-ing* form of the lexical verb), VVI (the base form of the lexical verb as an infinitive), VVN (the past participle form of the lexical verb), and VVZ (the *-s* form of the lexical verb).

As for the type of the syntactic chunks included in the analyses, I have counted both verbs occurring in main sentences, as well as verb forms occurring in subordinate clauses. When a relative clause was analysed, the *wh*-word was accounted for as a sign of a presence of a participant at its original position. For example, in sentence (1)

- (1) I saw the man *who* kicked the dog.

who is counted as the subject of the clause headed by *kicked*, and the semantics of this participant is seen as related to *man*. While in the example in (2)

- (2) I have seen the vase *which* they broke.

which is counted as the direct object of *break*, and the semantics of this participant is explored in relation with *vase*.

A similar approach has been used in analysing questions, as in the examples in (3) below, where the question word *whom* is counted as a sign for the presence of an effected participant, in this case an *Absorber*; while the question word *when* is counted as an element specifying the time at which the situation took place.

- (3) a. *Whom* did they shoot?
 b. *When* did they shoot him?

The corpus data for each verb has been analysed in several aspects either concerning the sentence as a whole or focusing on different parts of it.

- 1) USAGE - how the verb is used in every sentence
- 2) SUBJECT – what type of subject is allowed with that verb
- 3) DIRECT OBJECT – what type of direct object the verb specifies for (if any)
- 4) COMPLEMENTATION – what other lexically encoded participants and/or non-lexicalized elements may co-occur with this verb

The analyses of the data included both a syntactic and a semantic part. The first one reflects some noteworthy syntactic characteristic that may have consequences for the semantic representation of the verbs. It includes information about whether the verb is used in Active or Passive voice, whether a direct object is overtly realized or non-realized in the sentence, and when there are other participants whether they are syntactically realized as adverbs, prepositional phrases or sentential elements.

The second part reveals the semantic features associated with the participants, as well as the usage of the sentence as a whole. Therefore, I have distinguished between sentences where *Concrete Physical Impact* was implicated, and those with *Extended Meaning*. This difference is demonstrated in the pairs a) and b) in the examples in (4) to (6) below, where the a) sentences have a *Concrete Physical Impact* sense, while the b) sentences employ an extended meaning of the verb at hand.

- (4) a. ...he hit me hard on the face (BNC: FR6 54)
 b. The Recession has hit the Group in all major markets. (BNC: HBC 10)
- (5) a. ... their van was struck by a police car ... (BNC: CBF 14301)
 b. ... he was suddenly struck by inspiration (BNC: A7H 681)
- (6) a. He cast the paper aside impatiently. (BNC: G36 1118)
 b. ... she casts a heavy shadow over present negotiations... (BNC: AJ6 840)

The semantic values ascribed to the participants as present in the situations denoted by the set of sentences are introduced in the next section.

4.1.2 Terminology

The treatment of the semantic roles of the participants involved in a situation varies to a considerable degree among linguists not only across the different theoretical approaches but also within the particular theories. A recent online survey conducted by Murphy & Vogel (2006) aimed at evaluating the reliability of judgements of *thematic roles* (Dowty 1986, 1991, Jackendoff 1990) made by linguistic annotators. The overall results showed that the degree of agreement varies drastically for the different *roles*,¹ being highest for *Time* and *Agent* (95% and 94% mean majority agreement respectively) and lowest for *Instrument* (48%), *Percept* (49%), *Theme* (54%), and *Recipient* (59%).

Therefore, in the light of the current terminological diversity in linguistics, I must introduce first the terms which I have used for the semantic features that could be ascribed to the participants in a situation denoted by verbs. At the same time, for each of

¹ Although I do not agree with the application of the term *roles* here, I use it simply to present the results as they were discussed by the authors of the study (Murphy & Vogel, 2006). Thus, on their account, no distinction between a participant in the situation and a phrase specifying the event in some way is made, hence every single constituent in the sentence receives a *role*.

the terms I give illustrative examples taken from the corpus data analysed, thus introducing the verbs discussed in the research.

The core set of values assigned follows the formal description of the verbal sign given in detail in Chapter 3 (section 3.3) and are presented with examples in the sections to follow.

4.1.2.1 Source or Initiator

Whenever there is a *Force* emission in the situation denoted by the verb at hand, we have a participant releasing the *Force*, which is generally referred to as *Source* (cf. The Sign Model terminology (Dimitrova-Vulchanova 1996/99) and the discussion in the present work, section 3.3).

The *Source* can be any mechanism capable of discharging the energy needed to carry out the situation as illustrated in the examples in (7) below.

- (7) a. ... *a massive wall of water* smashed their caravans to matchsticks...
(BNC: CH2 7641)
- b. ... *silniat vjatiŕ sŭbori ogromno dŭrvo* ... (LCWB²)
strong-the wind knocked enormous tree
- b' *The rough wind* knocked an enormous tree.
- c. *Vŭlnite hvŭrljaha prŭski* ... (LCWB)
waves-the threw splashes
- c' *The waves* threw splashes.

In addition, if the participant releasing the force controls the situation, it can be further specified as *Initiator*, as already discussed in section 3.3.3.4. Thus, in the examples in (8) the italicized participants are regarded as *Initiators*.

² Since the Large Corpus of Written Bulgarian (LCWB) is not searchable over the internet and the references are not abbreviated as in the BNC, I do not include the specific text reference. Hence I indicate only that the example is taken from this corpus by using the abbreviation LCWB in parenthesis.

- (8) a. ... *they* begin to smash small bottles of some holy liquid ... (BNC: FP1 1561)
- b. ... *he* thought Sir Walter was on fire and threw a jug of ale over him. (BNC: G25 940)
- c. *He* scratched his head, embarrassed. (BNC: AEB 1937)

Whenever a participant with the value of *Source/Initiator* is explicitly mentioned in the sentence, it either appears in the subject position in active sentences as in the examples in (7) and (8) above or in a *by*-phrase if the sentence is in passive as in the examples in (9) below.

- (9) a. ... he was hauled out of the room *by his Aunt Janice* ... (BNC: GOA 2307)
- b. Muž bil udaren *ot kostenurka*, padnala ot nebeto. (LCWB)
man been hit by tortoise fallen from sky-the
- b' A man was hit *by a tortoise*, which had fallen from the sky.

This holds even if the verb is used in sentences with *Extended Meaning* as illustrated by the examples in (10) below.

- (10) a. I was struck *by his attitude*. (BNC: A6E 202)
- b. ... who are now being hit hardest *by the recession*. (BNC: ABS 2031)
- c. ... another pause slashed and stabbed *by the insect life*. (BNC: FAJ 1587)

4.1.2.2 Source Extension

When a participant with the value of *Source* performs by means of an instrument or part of its body is used as a mediator in releasing the force, the element denoting this instrument/body part is referred to as *Source Extension*. Usually it is syntactically

realized with a prepositional phrase as in the examples in (11) below, but it can also appear in the subject position as illustrated in the subsequent examples in (12).

- (11) a. If she hit me *with a stick*, ... (BNC: FR6 445)
 b. Tunney stabbed towards the door *with a thumb*. (BNC: GVL 3025)
 c. Mistür Kijn se izkiska i plesna *s dlan* po müršavoto si bedro. (LCWB)
 mr Keen refl.cl. giggled and slapped with palm on lean-def his thigh
 c' Mr Keen giggled and slapped his palm at his lean thigh.
- (12) a. *Lumberjack's paw* tapped the floor. (BNC: C86 2415)
 b. ... *the arrow* hit him in the eye. (BNC: HXF1263)
 c. ... *izvednůž nečija dlan* go pljasna lekičko po djasnoto ramo. (LCWB)
 suddenly someone's palm him slapped lightly on right-the shoulder
 c' Suddenly *someone's palm* slapped him gently on the right shoulder.

The participant with the value of *Source Extension* can also appear in the direct object position as envisaged in the Sign Model and illustrated by the English examples in (13) and the Bulgarian examples in (14) below.

- (13) a. Lars tapped *his fingers* reflexively against his wax tablet ... (BNC: G1M 667)
 b. ... slapping *his hand* on his knees ... (BNC: FP6 346)
 c. ...we can also raise our arm and stab *our finger* in the air towards it.
 (BNC: ADF 997)
- (14) a. Kalvin potupa *edin pribor* vůrxu kitkata si. (LCWB)
 Calvin tapped one tool on/against wrist-the refl.poss.
 a' Calvin tapped one tool against his wrist.
- b. ... plesna *kamšika* i ni povleče iz njuorskite ulici. (LCWB)
 slapped whip-the and us dragged along New York-poss.-the streets

- b' (He) slapped the whip and began dragging us along the streets of New York.
- c. Mušna prüst v otvora. (LCWB)
stuck finger in opening-the
- c' (He) stuck a finger in the opening.
- d. Rüğna pistoleta v gürdite mu. (LCWB)
stabbed gun-the in chest his
- d' (He) thrust the gun in his chest.

The overt realizations of the participant with the value of *Source Extension* in the examples in (11), (12) and (13) are viable mappings to syntax predicted by the theory, as already discussed in Chapter 3 (section 3.3). Given that *Source Extension* is that part of the *Source*, which directly acts upon the element absorbing the force emission it is a) *Criterion* by contact, which is reflected in its realization in the subject position, and b) an element with the value of *Monodeveloper* that marks it as a possible candidate for the direct object position.

4.1.2.3 Absorber or Limit

The participant, on which the *Force* is applied, is referred to as *Absorber* or *Limit*. The distinction between *Limit* and *Absorber* has been introduced in section 3.3.3.1 and now it is again demonstrated with the difference between the pairs of sentences a) and b) in all the examples from (15) through (17). The italicized items in the a) sentences are regarded as *Limit* while the ones in the b) sentences are counted as *Absorber*.

- (15) a. ... and kicked *him* in the face ... (BNC: CH2 7047)
b. ... kicks *his football* into the neighbours' garden... (BNC: CHR 273)
- (16) a. ... and an old crony to slap *me* on the back. (BNC: ECU 499)

- b. Daniel slapped down *a florin* on the counter. (BNC: EA5 1196)
- (17) a. He scratched *his moustache* thoughtfully. (BNC: FSR 1590)
 b. Don't tap the screen, you'll scratch *it*. (BNC: KD5 5918)

Thus, the sentences in (15b), (16b) and (17b) illustrate the possibility of the *Absorber* to undergo a movement or a process as result of the force applied to it. Whereas, *Limit* is regarded as the last entity in the *Force* chain, and no such process or movement obtains as can be seen in the examples (15a), (16a), and (17a).

Especially subtle is the distinction between the two examples in (17), where the difference involving the two senses of *scratch* is observed. Whereas in the use of *scratch* in (16a) only contact is implicated, in the situation described in (16b) the participant in the direct object position undergoes a monodevelopment with *Medium: Integrity*. The difference is even more apparent if the verb is used in the middle construction as shown in the examples in (18) below.

- (18) a. **His moustache* scratches easily.
 b. This screen scratches easily.

This observation is in line with recent research conducted by Hare et al. (2003) (presented in section 3.2 above) where the authors argue that the meaning of verbs is at least one of the determinants of verbs subcategorization preferences.

The distinction between *Absorber* and *Limit* is further discussed in section 4.2, where I also focus on the diverse patterns of syntactic behaviour linked to the difference in the semantics of these participants.

4.1.2.4 Place of Contact

The *Place of Contact* value has been assigned to those elements, which semantically are part of the participant that bears the value of *Limit*. Syntactically, they usually occur in a prepositional phrase in those sentences where *Limit* is already explicitly present, either

in the direct object position, as in the examples (19) and (20) for English and Bulgarian respectively, or in the subject position of a passive construction, as in the examples in the English examples in (21a) and (21b) and the Bulgarian example in (21c). The example in (21d) illustrates a case of the Bulgarian *se*-passive.

- (19) a. ... a man tapped her *on the shoulder* ... (BNC: GU9 443)
 b. Dr Frome slapped himself violently *on the thigh*. (BNC: H8Y 2962)
- (20) a. Njakoj rŭgaše Ivančev *v rebrata*. (LCWB)
 someone stabbed Ivanchev in ribs-the
 a' Someone was poking Ivanchev in the ribs.
- b. Tja silno go udari po glavata. (LCWB)
 she violently him hit on head-the
 b' She violently hit him on the head.
- (21) a. He had been stabbed *in the chest*. (BNC: CBF 14434)
 b. Angela was kicked *in the abdomen and on the chin* as the villagers rushed to help. (BNC: CH2 7047)
 c. ... tja be probodena *v gŭrba*. (LCWB)
 she was stabbed in back-the
 c' She was stabbed *in the back*.
- d. Toj naistina se plesna *po čeloto*. (LCWB)
 he really refl.cl. slapped on forehead-the
 d' He really slapped himself on the forehead.

As we can see in (21b) the *Place of Contact* element may be repeated thus indicating an iteration in the aspectual specification of the verb as denoting the situation at hand.

However, an element with the value of *Place of Contact* may also be promoted to the position of the direct object if *Limit* is not overtly present as demonstrated in the examples in (22) and (23) for English and Bulgarian, respectively.

- (22) a. It made me want to slap *his face*. (BNC: G07 2061)
 b. The Doctor tapped *the side of his nose*. (BNC: G1M 944)
- (23) a. Probodox *sŭrceto mu* dokato se otbranjavax. (LCWB)
 stabbed heart-the his while refl.cl. defended-1p.sg
 a' I stabbed *his heart* while I was defending myself.
- b. ... potupa koljanoto mu prosto ot priatelsko čuvstvo. (LCWB)
 ... tapped knee-the his simply from friendly feeling
 b' (He) tapped his knee just out of a friendly feeling.

Although relatively rarely encountered in the analyzed corpus data, the promotion of the *Place of Contact* element as the subject in passive constructions is also a valid mapping to syntax, which was again predicted in the Sign Model. Similarly to the previous case, the constraint is that no participant with the value of *Limit* is overtly realized.

- (24) a. ... useštax bolka v *sŭrceto si*, sjakaš beše probodeno ot ostra,
 felt-1p.sg pain in heart-the refl.poss.cl., as if was stabbed from sharp
 otrovna igla. (LCWB)
 poisonous needle
 a' I felt pain in my heart as if *it* was stabbed by a sharp poisoned needle.
- b. Seruyange said that he was beaten on the eye, chest, and *his leg* was stabbed. (*The Monitor* online)
- c. But suddenly, the light went off, and *Alex's leg* was stabbed with knife.

One can argue, however, that these cases in fact prove that *Place of Contact* should not be counted as a separate participant, since in all the examples in (22) to (24) the same element may also be regarded as *Limit*. Although that might be the primary intuition, there are many examples in the corpus data leading to the conclusion that this element has indeed a special status and should not be equalled to *Limit*. For example, compare the sentences in (25) below.

- (25) a. ... a man tapped her *on the shoulder* ... (BNC: GU9 443)
 b. A man tapped her shoulder.
 c. She dabbed his forehead. (BNC: AC2 845)
 d. *She dabbed him on his forehead.

The ungrammaticality of (25d) compared to the grammaticality of (25a) illustrates the difference in the status of the two PP_{on} phrases. While *on the shoulder* in (25a) can be ascribed the value of *Place of Contact*, *his forehead* in (25c) is clearly a *Limit* as it is seen by the impossibility to include another participant (*him*) that would claim the value of *Limit* in the sentence in (25d). Therefore, I have distinguished between the two of them and I regarded *Place of Contact* as a separate value.

4.1.2.5 Path values

Many sentences contained one or more phrases specifying a *Path* component.³ Based on the empirical data I have been able to distinguish four different values with respect to *Path*. The value of *Origin* is ascribed to a phrase which denotes the starting point of the movement of the Absorber (hence also Mover) in result of the force emission. Examples of *Path Origin* phrases are presented in (26) below.

³ The terminology used in this work generally follows the terminology used in recent work on Motion encoding by Dimitrova-Vulchanova (cf. Dimitrova-Vulchanova 2004, among other papers).

- (26) a. ... the truck drivers pushed them *out of the moving cab*. (HH0 1533)
 b. V tozi moment pŕvijat go butnal *ot prozoreca* ... (LCWB)
 in this moment first-the him pushed from window-the
 b' At this moment, the first one pushed him *from the window*.

The *End of Path* value is assigned to those elements that mark the end point of the path of the Mover as illustrated by the examples in (27) below.

- (27) a. ... the contracting parties had cast lumps of irons *into the see*, ...
 (BNC: G3C 1263)
 b. Posle metna klečkata v *kartonenata čaška*. (LCWB)
 then cast stick-the in cardboard-the cup
 b' Then he cast the stick *into the cardboard cup*.

Beside the points of origin and/or end, a certain region of the path may be specified and in this case, the value of *Path Length* has been assigned. The PPs in italics in the sentences in (28) below can serve as examples of such cases.

- (28) a. ...and watched Tom drag his small mattress *past the window*. (BNC:
 CAB 447)
 b. Eliūt bukvalno ja vlačí *prez ostavaštoto razstojanie do lagera*.
 (LCWR)
 Elliot literally her dragged through remaining distance to camp-the
 b' Elliot literally dragged her *over the remaining distance to the camp*.

Finally, there were also instances where only the direction of the movement was stated without any indication of the starting or the ending point of the Mover. Therefore, the value of *Path Orientation* was introduced to account for the semantic status of these elements.

- (29) a. Maggie lunged out an arm and hauled the girl *towards her*. (AN7
1248)
- b. Džonti otstūpi nazad i ritna kufara *vstrani*. (LCWB)
Jaunty stepped back and kicked suitcase-the aside
- b' Jaunty stepped back and kicked the suitcase *aside*.

The four values described above may occur alone as already illustrated above or in various combinations within the single clause as demonstrated in the sentences in (30) below.

- (30) a. Then he dragged himself *through the crowds* to the quiet cranny of the Ibis Boat Club at Chiswick. (BNC: AHU 838) (*Length + End*)
- b. She began hauling suitcases *out of the hall*, onto the grass. (BNC: APM 48) (*Origin + End*)
- c. Toj se vlačēše *napred* prez xrastalacite. (*Orientation + Length*)
he refl.cl. dragged forward through bushes-the
- c' He was dragging himself forwards through the bushes.
- d. Toj metna saksijata *ot/prez prozoreca* kūm minavaštija otdolu sused.
he cast flowerpot-the from/through window-the towards passing-the
beneath neighbour (*Origin/Length + Orientation*)
- d' He cast the flowerpot from/through the window towards the neighbour
that was passing beneath.

4.1.2.6 Other terms used in the analyses

Much in the tendency of the traditional θ -roles, the value of *Beneficiary* has been assigned to the entity for which the action was performed. Those were very rare in the corpus data analysed.

(31) ... if the food is cut up *for the patient*. (BNC: AS0 254)

Resultatives are those elements that describe a state, which is a result from the process, denoted by the verb. Thus, not only adjectives were considered as the examples in (32a-32c), but also prepositional phrases as show in the sentences in (32d) and (32e).

- (32) a. ... he shot *dead* a council worker... (BNC: K1K 548)
 b. ... the Blackshirts had smashed *open* his tailor's shop. (BNC: CHG 312)
 c. She pulled the zip of her bag *shut* ... (BNC: JXT 2461)
 d. ... Philip Shehadi, was found stabbed *to death* ... (BNC: HL5 2341)
 e. Katie cut hers up *into five hundred pieces*. (BNC: KCK 1274)

The elements that describe the way in which a situation takes place were given the value of **Manner**. Sentences containing various *Manner* phrases are given in the English examples in (33) and the Bulgarian examples in (34) below.

- (33) a. Melissa dabbed *daintily* at her mouth. (BNC: HGT 2716)
 b. ... it hit the floor *with a crack*. (BNC: AD9 825)

- (34) a. ... vŭlcite *bavno* se vlačexa *edin zad drug*. (LCWB)
 wolves-the slowly refl.cl dragged one behind another
 a' The wolves were crawling slowly one after another.
 b. ... Ksintija *taka* go blŭsna, *če toj politna v ŭgŭla* ... (LCWB)
 Ksintia thus him pushed that he flew in corner-the
 b' Ksintia pushed him *so hard that he flew into the corner*.

The **Locative** value has been ascribed to those entities that specified the location at which the situation as a whole takes place, as exemplified in (35) below.

- (35) a. ... a jeweler was stabbed *in his shop* ... (BNC: K1R 1835)
 b. Mladež strelja po štaba na "Aum" v *Tokio*.
 youth shot at headquarters-the of Aum in Tokyo
 b' A young man shot at the Aum headquarters in Tokyo.
- c. ... Toj bil blūsnať ot avtomobil *na pūt za trenirovka*. (LCWB)
 he been hit by car on way to training
 c' He has been hit by a car *on his way to the training*.

The phrases indicating the time at which the situation takes place received the value of *Time* as illustrated in (36) below.

- (36) a. ... he was shot down *after ten minutes in the air*. (BNC: CA8 423)
 b. ... and scratched his fuzzy head, *as he sat on his cart* ... (BNC: EA5 991)
 c. *Sled tova* toj ja mušna v džoba si ... (LCWB)
 after that he her put in pocket-the his
 c' *After that*, he put it in his pocket ...
- d. Gabi potupa Psalterij po xūlboka, dokato Kris minavaše pokraj tjax.

I have used the term *Contingency* (Quirk et al. 1985) for all the phrases that imply logical relationship between two events. Therefore, it comprises among others the values of *Reason* (the situation that motivates the one analyzed), *Condition* (the situation that would bring about the one analyzed) and *Purpose* (the aim of the situation) as illustrated in the examples in (37) below.

- (37) a. They nearly came to scratching each other *because Janice says they've only got an old banger*. (BNC: HWE 1009) (*Reason*)
 b. ... those who push themselves forward into police investigations *unless directly called upon*. (BNC: CDB 1991) (*Condition*)

- c. ...he'd scratch them to pieces *if they didn't obey him*. (BNC: EWC 1315) (Condition)
- d. Potupa džoba na pantalona si, *za da se uveri, če*
tapped pocket-the of trousers-the poss.refl.cl., for to refl.cl. assure, that
signalnijat pistol e vŭtre ... (LCWB) (Purpose)
signal-the pistol is inside
- d' He tapped the pocket of his trousers to make sure the signal pistol is inside.

The terms *Quantification* and *Measure* were used to signify phrases that denote repetition or a certain extent associated with the situation as exemplified in (38).

- (38) a. Jakub struck *again*. (BNC: HTH 3254)
- b. The teacher used a stick to *repeatedly* tap or prod pupils ... (BNC: GUR 289)
- c. ... then dabbed at Dot's cheeks *till they were sore*. (BNC: AC5 731)
- d. Njakoј butna vratata *samo kolkoto da ustanovi, če e zaključena*.
someone pushed door-the only just to find that is locked
- d' Someone pushed the door *just enough to find out that it was locked*.

Elements affecting the truth-value of the sentence or logically modifying its statement were regarded as expressing *Modality*.

- (39) Korab *edva* ne se blŭsnał v tjaloto.
ship almost not refl.cl.. hit in body-the
A ship almost hit against the body.

Although specified above, most of the values presented in section 4.1.2.6 are not considered in the actual analyses of the data to the extent that they are not included in the lexical representation of the verbs at hand. Thus, their presence was registered only in the tables with the rest of the results (cf. Appendices B and C for the tables of results in English and Bulgarian, respectively).

4.2 Verb Grouping

As mentioned earlier in section 2.1, the verbs investigated in this project were classified by Levin (Levin, 1993) according to the various alternations in which they can participate. On this classification, however, every verb belongs to several classes, where *Verbs of Contact by Impact*, *Verbs of Throwing*, *Verbs of Exerting Force*, *Verbs of Sending and Carrying*, *Verbs of Obtaining*, *Verbs of Separating and Disassembling*, and *Verbs of Psychological State* are only some of the possibilities. Therefore, I have aimed at a more detailed analysis of the semantics of these verbs with respect to the mapping from conceptual structure to syntactic realization of participants as reflected in the corpora. The results from this investigation for English and Bulgarian verbs can be seen in Appendices B and C, respectively.

Thus, in a more elaborate assessment, the examined verbs manifested a set of semantic features shared amongst them, namely, the presence of *Force emission* and obligatory physical contact of the *Source/Source extension* with the *Absorber/Limit*. Based on this set of features, we can generally distinguish two major types of situations lexicalized by the verbs at hand – situations that focus on the contact attained as result of a force emission and situations where the focus is set on a conditioned event, which either entails the contact or results from it. This grouping mirrors the humans' perception of the world and the way situations are chunked into sub-events that may differ cross-linguistically. Hence, the semantic roles of the participants involved in the situation the verb at hand lexicalizes naturally reflect the variety of alternations common to the verbs in each of the groups. In addition, some verbs can lexicalize events from both types as they share features with both groups.

4.2.1 Verbs of Contact

As briefly mentioned above, the verbs in this first group denote a situation that focuses on the *Contact* attained as the result of a *Force* emission. Therefore, the main participants in the lexicalized event are the *Source* (the one that releases the *Force*) and the *Limit* (the participant that absorbs the force as the last entity in the Force chain). A

Source Extension (the part of the source or the instrument used to achieve contact with the *Limit*) and a *Place of Contact* (the part of the *Limit* identified as the exact place where contact was attained) may further specify the situation.

The central members of this group are the English verbs *hit*, *smash* (where only contact implied), *tap*, *slap*, *kick*, *strike*, and *stab*, as well as their Bulgarian correlates *udarja*, *capna*, *potupam*, *pljasna*, *ritna*, *ucelja*, and *proboda*, respectively.

The basic cell of this group of verbs has the abstract representation given in Fig.1 below (with some variation for the different verbs).

Aspectual specification: +2-point		
Element specification:		
<u>Constituency</u>	<u>Force</u>	<u>Monodevelopment</u>
[+Contact with 3]	Source ₁	
[Mediator] ₂	Source extension ₂	Mover ₂
		Monodevelopment _a
		<i>Phasing</i> : +2-point
		<i>Medium</i> : Location
		Line of Trajectory
		End: Contiguous to 3
[+Contact]	Limit ₃	
[part of Limit ₃]	(Place of Contact ₄)	

Fig.1 Abstract lexical representation of Verbs of Contact

Since *Contact* is the criterial aspect here, the event is oriented towards achieving this contact and is exhausted when the contact is attained. Therefore, all the participants that are specified for [+*Contact*] will count as *Criterial* in this type of situation. Consider the grammaticality of the following examples, both in Bulgarian and in English, respectively.

- (40) a. *Toj udari.
 he hit
 a' *He hit.⁴
- b. Toj udari *(stenata) s jumruk.
 he hit wall-the with fist
 b' He hit *(the wall) with his fist.
- c. Toj se udari (na pregradata pred kaminata).
 he refl.cl. hit on fence-the in front of fireplace-the
 c' ...he hit *(himself) on the fender... (BNC: HWM 756)
- d. Topkata se udari *(v stenata).
 ball-the refl.cl. hit in wall-the
 d' The ball hit *(the wall).

Both the Bulgarian and the English examples in (40a-a') and (40b-b') are unacceptable without the direct object, as it is the overt expression of a *Criterion* element. However, (40c) is grammatical with and without the prepositional phrase, yet on different readings. This is due to the properties of the Bulgarian reflexive *se*-constructions (*se* is viewed as a contracted version of *sebe si (oneself)*, which was already discussed in Chapter 1). Therefore, the Bulgarian sentence in (40c) can be grammatical even without the prepositional phrase and its meaning is equivalent to the English *he hit himself*. In English, however, the reflexive must be overtly expressed, as seen in the unacceptability of (40c') given that *himself* is omitted.

In addition, the unacceptability of (40d) without the prepositional phrase, demonstrates that indeed an overtly expressed phrase specifying the *Limit* must be

⁴ I will not discuss cases of apparent object omission in generic phrases such as 'hit and/or miss' and 'hit and run' or in sentences like the examples (1) and (2) below.

(1) The craft of boxing is to hit and not be hit,'; responds Eubank.(BNC: ACP 2221)

(2) If your second shot hits you can fire again. (BNC: CN1 538)

present in all the cases where the *se*-construction is not reflexive, e.g. it cannot be paraphrased as *hit oneself* as the ball cannot hit itself. Instead, the *se*-construction in (40d) is considered an Absolutive *se*-construction (cf. Dimitrova-Vulchanova (1996/99) and the discussion in Chapter 1 of the present work.)

Although the focus of the events lexicalized by the first group of verbs is set on the contact itself (where no monodevelopment is included), a participant with the value of *Monodeveloper* may also be realized syntactically, if it acts as a *Mediator* between the *Source* and the *Limit*. Thus, the default *Monodeveloper* is usually the *Source Extension*, which performs a monotonic development with *Medium:Location* and its movement is said to be Contiguous to 3 (Fig.1 above), i.e. it attains contact with the *Limit*. Therefore, it counts as a criterial element and can appear in a subject position. This was envisaged by The Sign Model, since it was predicted that the *Source* does not need to act by itself, but may use a *Mediator*, which can be overtly realized as the subject of the headed sentence. This alternation is further referred to as the *Instrument Subject Promotion*⁵ (in Levin (1993) *Instrument Subject Alternation*, p. 80) and it is illustrated in the examples in (41) below.

- (41) a. Toj udari masata s rŭka.
 he hit table-the with hand
 a' He hit the table with his hand.
- b. Rŭkata mu udari masata.
 hand-the his hit table-the
 b' His hand hit the table.

Another alternation explainable in terms of *Criteriality* is dubbed here the *Place of Contact Promotion* (on Levin's classification this is the *Body-Part Possessor Ascension Alternation* (Levin 1993, p. 71.) That is, whenever a participant with the

⁵ Where appropriate, the terminology follows the names given by Levin (Levin 1993). In cases where a new name is introduced for an alternation also mentioned in Levin (1993), a reference to the name given by Levin is also included in brackets.

value of *Place of Contact* is specified in a prepositional phrase, it can also be promoted to the direct object position, as it is a part of the *Limit* and thus, is *Criterial* as 'an item in a contact relation' (Dimitrova-Vulchanova 1996/99).

- (42) a. Udarix ja s kamük *po glavata*.
 hit-1p.sg. her-Acc.cl. with stone on head-the
 a' I hit her *on the head* with a stone.
- b. Udarix *glavata i* s kamük.
 hit head-the her-poss.cl.
 b' I hit *her head* with a stone.
- c. Udarix i *glavata* s kamük.
 hit her-Dat.cl. head-the with stone
 c' I hit *her head* with a stone.

The Bulgarian examples in (42b) and (42c) differ only syntactically (on the linear realization). However, they are semantically equivalent and in both sentences, the direct object position is taken by a participant marked with the value *Place of Contact*.

Since each of the criterial participants in the situation may be realized independently, we can observe different combinatorial patterns. Thus, for example, the so-called *Instrument Subject Promotion* and the *Place of Contact Promotion* may appear simultaneously as illustrated in the examples in (43).

- (43) a. Kamüküt ja udari *po glavata*.
 stone-the her- Acc.cl. hit on head-the
 a' The stone hit her on the head.
- b. Kamüküt udari *glavata i*.
 stone-the hit head-the her-poss.cl.
 b' The stone hit her head.

- c. Kamüküt i udari *glavata*.
stone-the her-Dat.cl. hit head-the
- c' The stone hit her head.

Again, we have the syntactic variation between (43b) and (43c) in Bulgarian, as already discussed for the examples in (42).

We can also identify a type of *Place of Contact Promotion* in passive constructions where the referent of the subject in the sentence bears the *Place of Contact* value. This syntactic realization is predicted by The Sign Model, since this participant is marked as *Criterial* and thus it is promoted to the position of a subject in passive constructions, whenever the *Limit* is not overtly realized as shown in the examples in (44a) and (44b) below.

- (44) a. Tja beše udarena s kamük *po glavata*.
she was hit with stone on head-the
a' She was hit on the head with a stone.
- b. *Glavata i* beše udarena s kamük.
head-the her-poss.cl. was hit with stone
b' Her head was hit with a stone.

As already mentioned earlier in the analysis of the example in (40c) and (40d) above, if a member of the *Verbs of Contact* group participates in a *se*-construction in Bulgarian, it will either entail a reflexive meaning or else the clause will be regarded as an Absolute *se*-construction. The preferred reading of the sentence depends on the semantic features of the participant realized as the subject.

Thus, if the referent of the subject is marked with the value of *Source*, the preferred reading would be the reflexive. Therefore, the meaning of the sentence in (45a) will be generally the same as the meanings of the sentences in (45b) and (45c) below.

- (45) a. Toj se udari po kraka.
 he refl.cl. hit on leg-the
 a' He hit himself on the leg.
- b. Toj udari kraka si.
 he hit leg-the refl poss.cl.
 b' He_i hit his_i leg.
- c. Toj si udari kraka.
 he refl.poss.cl. hit leg-the
 c' He_i hit his_i leg.

This is the default reading with the *Verbs of Contact* group and examples of reflexive reading of the *se*-construction with other verbs from the group are presented in (46) below.

- (46) a. Dr. Froum se pljasna silno po bedroto.
 dr Froum refl.cl. slapped violently on thigh-the
 a' Dr Frome slapped himself violently on the thigh. (BNC: H8Y 2962)
- b. Toj se capna po kraka.
 he refl.cl. smashed on leg-the
 b' He smacked his leg.
- c. Toj se potupa po kraka.
 he refl.cl. tapped on leg-the
 c' He tapped himself on the leg.

However, if the referent of the subject does not bear the value of *Source*, the *se*-construction would be regarded as an Absolutive *se*-construction as in the example in (40d), repeated also in (47) below.

- (47) Topkata se udari v stenata.
 ball-the refl.cl. hit in wall-the
 The ball hit the wall.

Yet, this is not a characteristic feature of the group, but is restricted only to the verb *udarja* (*hit*) (cf. the discussion in section 4.2.3).

So far, with the exception of the example in (47) above, I have discussed only cases, where the subject position in active sentences is occupied by participants marked with the values of *Source* or *Source extension*. However, the analyses of the available corpus data revealed also cases, where the referent of the subject was not marked by any of these values as illustrated in the examples in (48).

- (48) a. *Knigata* pljasna na zemjata.
 book-the slapped on ground-the
 a' The book slapped on the ground.
- b. *Topkata* capna vŭv vodata.
 ball-the smashed in water
 b' The ball fell in the water.

However, in the sentences in (48), the verbs at hand are used to denote a different type of situation and are therefore discussed separately (cf. section 4.2.3).

Some of the verbs, lexicalizing a situation of contact, may also specify for a different *End of the Trajectory Line*, which is reflected in the different preposition used when the *Place of Contact* element is introduced, as demonstrated in (49) below.

- (49) a. Toj se probode/bodna v kraka.
 he refl.cl. stabbed/pricked in leg-the
 a' He stabbed/pricked himself in the leg.

Although not always, the preposition *v* (*in*) generally indicates that the *Endpoint* of the trajectory of the *Mediator* (*Source Extension*) is inside the *Limit*. Even

though the *Mediator* is lexically encoded (cf. Fig.1 in the beginning of section 4.2.1), it does not need to be syntactically realized as illustrated in the example in (49) above. However, it cannot be eliminated and it is always conceptually present in the situation denoted by the verb at hand.

Verbs that do specify for an inner *End-point* are special in that they might intuitively be confused with *Verbs of Change of Integrity* (for example *cut*, cf. section 4.2.2.2 for further discussion). However, the syntactic patterns displayed in the corpora data demonstrate that this is not the case. Consider, for example, the difference in the grammaticality of the sentences in (50) below.

- (50) a. Chicken meat cuts easily.
 b. *Chicken meat stabs easily.

While the example in (50a) is completely acceptable, the example in (50b) is ungrammatical. In addition, when I asked native speakers to complete sentences (in a preliminary study), the completions I received for "*Chicken meat stabs ____*" (one of them was, for example, "*my gums when frozen*") indicated that the referent of the subject in this case cannot be assigned the value of *Monodeveloper*, but is perceived as a *Source*, instead.

Thus, it becomes evident that *stab (proboda)* does not belong in the same group with *cut (reža)*. Although a change of integrity may be conceptualized in the situation denoted by *stab*, this is obviously not lexicalized by the verb at hand.

4.2.2 Verbs denoting a Conditioned event

The situation denoted by the second group of verbs is a *Conditioned* event with two main participants *Source* and *Absorber*. The *Source*, in the default case, is co-indexed with the value of *Conditioner*. Moreover, the participant with the value of *Source* has control over the *Force* emission and the subsequent movement of the *Absorber* induced by that *Force*. Such a participant, characterized by the set of values [*Conditioner*, +*Control*], is also referred to as *Initiator* (cf. Dimitrova-Vulchanova 1996/99). As a

result of the absorbed force or energy, the element with the value of *Absorber* performs a *Monodevelopment* – a monotonic movement in a certain direction, and is thus specified as a *Monodeveloper* of type *Mover*.

Falling into this group are the English verbs *drag*, *pull*, *haul*, *push*, *scratch*, *cut*, *throw*, *cast*, and *shoot*, and their Bulgarian counterparts – *vlača*, *dŭrpam*, *teglja*, *butam*, *draskam*, *reža*, *xvŭrljam*, *mjatam*, and *streljam*. According to the aspectual specification of the verbs, the type of contact attained, and the *Medium* with respect of which the *Monodevelopment* is performed, this group can be further divided into three sub-groups discussed in sections 4.2.2.1, 4.2.2.2, and 4.2.2.3, respectively.

4.2.2.1 Verbs of Sustained Contact

The first sub-group contains verbs encoding a sustained contact as the English *drag*, *haul*, *pull*, and *push* and their corresponding Bulgarian verbs *vlača*, *teglja*, *dŭrpam*, and *butam*. These verbs differ in that the aspectual specification of the lexicalized event is [+protracted] and there is a specified Constituency component [+sustained contact].

Since both the *Source* and the *Absorber* are marked with the feature [+sustained contact], they are both criterial as each of them is 'an item in a contact relation' (Dimitrova-Vulchanova 1996/99). Thus, the basic cell of the verbs from this first sub-group appears to have the semantic representation given in Fig. 2 below.

Aspectual specification: + protracted			
Element specification:			
<i>Conditioning</i>	<i>Constituency</i>	<i>Force</i>	<i>Monodevelopment</i>
Conditioner ₁	[+ sustained contact]	Source ₁	
	[+ sustained contact]	Absorber ₂	Mover ₂
Conditioned			
Cell:			
Aspectual specification: +protracted			
Element specification:			
Monodevelopment _a			
<i>Element: 2</i>			
<i>Phasing: Protracted</i>			
<i>Medium: Location</i>			
<i>Path</i>			

Fig.2 Basic *Cell* of verbs lexicalizing a *Conditioned* event with sustained contact and a *Monodevelopment* with *Medium: Location*

The canonical realizations of the two main participants (marked with the indices 1 and 2 in Fig.2 above) are respectively the subject and the direct object positions in active sentences, as illustrated in the example in (51a). In passive constructions, however, the participant with the set of values [*Absorber, Mover*] is realized as the subject of the sentence and the participant with the set of values [*Conditioner, Source*] may appear in an optional *by*-phrase. This is demonstrated in the example in (51b) below.

- (51) a. His friends pushed the car to the top of the hill.
 b. The car was pushed to the top of the hill (*by his friends*).

In addition, an element with the value of *Path* (*origin, end, length, or orientation*) is also encoded in the lexical representation of the verbs at hand, specifying the trajectory of the movement performed by the *Absorber*. Although not syntactically obligatory, it is frequently realized overtly as an adverbial or a prepositional phrase, as illustrated in the English examples in (51) above and the Bulgarian examples in (52).

- (52) a. Tja go vlačeše *nadolu po ulicata*. (LCWB)
 she him dragged down on street-the
 a' She dragged him down the street.
- b. Vojnicite se dūrpat ot nego ... (LCWB)
 soldiers-the refl.cl. pull from him
 b' The soldiers are pulling back from him ...

Due to the nature of the lexicalized situation (including a *Monodevelopment* with *Medium: Location*), when a verb from this group participates in the Bulgarian *se*-construction, it behaves similarly to the *Manner of Motion* verbs – the referent of the subject performs a movement in a certain direction in a manner that matches the causative process denoted by the verb in its transitive form. This is illustrated with the examples in (53) below.

- (53) a. Momčeto se vlačeše sled roditelite si.
 boy-the refl.cl. dragged behind parents-the refl.poss.cl
 a' The child was dragging himself behind his parents.
- b. Deteto se tegleše nazad [daleče ot vodata].
 child-the refl.cl. hauled away [far from water-the]
 b' The child hauled himself away from the water.
- c. Momičeto se dūrpaše nazad.
 girl-the refl.cl. pulled backwards
 c' The girl was pulling (herself) backwards.

- d. Skitnikūt se butaše napred v tūlpata.
 wanderer-the refl.cl. pushed ahead in crowd-the
 d' The wanderer pushed his way ahead in the crowd.

Reflexivization of this type of verbs may be regarded as a means in some natural languages to use transitive verbs in lexicalizing intransitive situations, since in all the sentences in (53), the reflexive clitic *se* cannot be substituted by its full counterpart *sebe si* ('oneself'), thus denoting a truly transitive event. Instead, the verbs at hand are used to denote motion in a corresponding manner, which is also confirmed in the example in (54) illustrating that the sentences in (53) may be paraphrased using a true motion verb and an Adverbial of Manner.

- (54) Momčeto vūrveše sled roditelite si, vlačejki se.
 boy-the walked behind parents-the refl.poss.cl dragging refl.cl.
 The child was walking behind his parents, dragging himself.

Although all of the examples in (53) describe a situation, where the referent of the subject in the sentence is defined as [*Conditioner*, +*Control*], we may also encounter cases where the subject position is taken by a participant that is not specified for any of these features. Moreover, this participant may in fact bear the value *Absorber*. Thus, a different perspective of the main situation is taken, as we have zoomed in on the *Conditioned* event only. This may be regarded as the limiting case, where no *Source* needs to be specified. Thus, the example in (55b) illustrates another instance of the Absolutive *se*-construction (cf. Dimitrova-Vulchanova 1996/99).

- (55) a. Momčeto vlačeše vūžeto (po zemjata).
 boy-the dragged rope-the on ground-the
 a' The boy dragged the rope (on the ground).
 b. Vūžeto se vlačeše *(po zemjata).
 rope-the refl.cl. dragged on ground-the
 b' The rope was dragging *(on the ground).

In this limiting case, however, the *Path* referent must be overtly realized or at least implicated (cf. the discussion in section 3.3.5.2), since it is also a participant in a contact relation and cannot be eliminated as seen in the illustrative example in (55b).

4.2.2.2 Verbs of Change of Integrity

The second sub-group of verbs denoting a conditioned event is dubbed *Verbs of Change of Integrity* and it includes the English verbs *scratch* and *cut* and their Bulgarian correlates *draskam* and *reža*. The main characteristic feature of the verbs falling in this group is the existence of a second *Monodevelopment* and the *Medium* with respect of which it is performed, as well as the [+2-point] aspectual specification.

Verbs of Change of Integrity lexicalize situations where, as a result of the force emission, two monodevelopments are induced. The first one (Monodevelopment_a in Fig.3 below) involves the participant marked with the set of values [*Source Extension*₂, *Mover*₂] and also specified as a sharp-edged tool, which undergoes change of position, i.e. it is a monodevelopment with *Medium*: Location.

The second monodevelopment (Monodevelopment_b in Fig.3 below) involves the participant with the set of values [*Absorber*₃, *Monodeveloper*₃], which takes part in a 2-stage process with *Medium*: Integrity, changing from [+intact] to [-intact] as an immediate result from Monodevelopment_a.

Thus, the basic cell of the *Verbs of Change of Integrity* must have the following semantic representation given in Fig.3 below.

Aspectual specification: +2-point			
Element specification:			
<u>Conditioning</u>	<u>Constituency</u>	<u>Force</u>	<u>Monodevelopment</u>
Conditioner ₁		Source ₁	
	[sharp-edged tool] ₂	Source Extension ₂	Mover ₂
		Absorber ₃	Monodeveloper ₃

Conditioned

Cell:

Aspectual specification: +2-point

Element specification:

Monodevelopment_a*Element: 2**Phasing: +2-point**Medium: Location*

Stage 1: Role: Out-of

Cell:

Aspect. Specification: -Dynamic

Element Specification: x_2

Stage 2: Role: Into

Cell:

Aspect. Specification: -Dynamic

Element Specification: x_2 in contact with x_3 Monodevelopment_b*Element: 3**Phasing: +2-point**Medium: Integrity*

Stage 1: Role: Out-of

Cell:

Aspect. Specification: -Dynamic

Element Specification: x_3 [+intact]

Stage 2: Role: Into

Cell:

Aspect. Specification: -Dynamic

Element Specification: x_3 [-intact]

Fig.3 Abstract Cell of Verbs of Change of Integrity

As already mentioned above, the presence of a second monodevelopment with *Medium: Integrity* is a distinctive feature of the verbs from this sub-group. The English verb *scratch*, however, may also lexicalize situations that end with the execution of the first monodevelopment, thus acting as a pure *Verb of Contact*. This ability of the verb *scratch* was illustrated also in section 4.1.2.3 where the difference between *Limit* and *Absorber* was demonstrated in the sentences in (17), repeated in (56) below.

- (56) a. He scratched *his moustache* thoughtfully. (BNC: FSR 1590)
 b. Don't tap the screen, you'll scratch *it*. (BNC: KD5 5918)

Whereas the situation in (56a) entails only that a contact is attained, e.g. the *Source Extension* is contiguous to the *Limit*, without affecting it, the situation in (56b) does entail a subsequent monodevelopment with *Medium: Integrity* on the part of the *Absorber*. Thus, we may distinguish two distinct senses of the English *scratch*, each used to lexicalize a different type of event. This ability of some verbs to denote situations from both types (either contact or conditioned event) is further discussed in section 4.2.3.

However, Bulgarian uses two separate items for the two situation types lexicalized by *scratch* in the examples in (56). Thus, the Bulgarian translations of *scratch* would be *češa* in the case of (56a) and *draskam* in (56b). Interesting as it may be, this situation was predicted by our hypothesis, since it was expected to find some cross-linguistic variation in parsing a situation and the lexicalization of its sub-events.

As this issue is addressed in detail in section 4.2.3 below, I will now return to the discussion of the corpus data for the present group of *Verbs of Change of Integrity*.

Considering the differences in the semantic representation of the verbs at hand (compare Fig.1 and Fig.2 with Fig.3 above), we did expect different patterns of syntactic behaviour than those already displayed by the verbs in the two groups discussed already.

- (57) a. Tja režeše mesoto.
 she cut-3p.sg-past meat-the
 a' She was cutting the meat.

- b. Mesoto se reže lesno.
meat-the refl.cl. cut-3p.sg easily
- b' The meat cuts easily.
- c. Tozi nož reže dori dūrvo.
this knife cut-3p.sg even wood
- c' This knife cuts even wood.
- (58) a. Toj draskaše s nokti (po) glatkata povūrnost.
he scratched with nails (on) smooth-the surface
- a' He was scratching the smooth surface with his nails.
- b. Zlatoto može da se draska s nokūt.
gold can to refl.cl. scratch with nail
- b' Gold scratches/can be scratched with a nail.
- c. Vsičko drugo draska bojata.
everything other scratches paint-the
- c' Everything else scratches the paint.

Although no participant with the value of *Source Extension* (in this case an *Instrument*) is syntactically realized in the sentences in (57a) and (57b), it is still part of the lexical representation of the verb. It cannot be eliminated because it is a *Criterion* element (as an item with the value *Monodeveloper* (in this case of type *Mover*) and an item characterised for inherent properties). The analysis presented here also conforms with the findings in Koenig et al. (2003) discussed earlier in Chapter 3. Thus, the existence of syntactic patterns like the ones in (57c) and (58c) is easily predicted.

In terms of Levin, this is the *Characteristic Property of Instrument Alternation* (Levin 1993, pp. 39-40), of which only the intransitive variant is explained with the characteristic property of the instrument used to accomplish the situation denoted by the verb. This account, however, does not explain the grammaticality of the sentences in (57c) and (58c), which are not intransitive. With respect to this, the element marked

with the set of values [*Source Extension*, *Mover*] behaves in exactly the same manner as the corresponding participant in the representation of the verbs from the first group – *Verbs of Contact*. This is apparent when we compare the examples in (41), given in section 4.2.1 above, with the sentences in (59) below.

- (59) a. Tja nadraska bojata na kolata *s telenata četka*.
 she scratched paint-the of car-the with wired-the brush
 a' She scratched the paint of the car with the wired brush.
- b. *Telenata četka* nadraska bojata na kolata.
 wired-the brush scratched paint-the of car-the
 b' The wired brush scratched the paint of the car.

Thus, it becomes evident that the *Characteristic Property of Instrument Alternation*, proposed by Levin, does not account for the syntactic pattern observed. Instead, it is suggested here that the set of semantic properties ascribed to the relevant participant must be used in predicting the possible syntactic positions, which it could occupy in the sentence.

Now compare the sentences in (59) with the examples in (60a) and (60b) below.

- (60) a. Tja nadraska bojata na kolata na telenata ograda.
 she scratched paint-the of car-the on wire-the fence
 a' She scratched the paint of the car on the wire fence.
- b. **Telenata ograda* nadraska bojata na kolata.
 wire-the fence scratched paint-the of car-the
 b' *The wire fence scratched the paint of the car.

While *wired brush* is an item with the value of *Source extension* in (59) and in addition, it is also a *Mover*, *wire fence* neither is marked as a *Source extension*, nor performs a *Monodevelopment* of any kind, which naturally results in the unacceptability of the sentence in (60b) as already predicted by the theory.

4.2.2.3 Verbs of Initial Contact and Ejection

The third sub-group of verbs denoting a conditioned event is named *Verbs of Initial Contact and Ejection* and it includes the English verbs *throw*, *cast*, and *shoot*, and the respective Bulgarian verbs *xvŭrljam*, *mjatam*, and *streljam*. Their aspectual specification is [+2-point], and their core property is the *Ejection* component in the beginning of the event (Monodevelopment_a in Fig.4 below), as well as the subsequent motion induced by the initial force emission (Monodevelopment_b in Fig.4 below).

Aspectual specification: +2-point

Element specification:

<u>Conditioning</u>	<u>Control</u>	<u>Force</u>	<u>Monodevelopment</u>
---------------------	----------------	--------------	------------------------

Conditioner₁

[+Control]₁

Source₁

Source Extension₂

Absorber₃

Mover₃

Conditioned

Cell:

Aspectual specification: +2-point

Element specification:

Monodevelopment_a

Stage 1: Role: Out-of

Cell: Aspect. Specification: -Dynamic

Element Specification: In-ness

Element: 3

Host₂

Stage 2: Role: Into

Cell: Aspect. Specification: +Protracted

Element Specification:

Monodevelopment_b

Element: 3

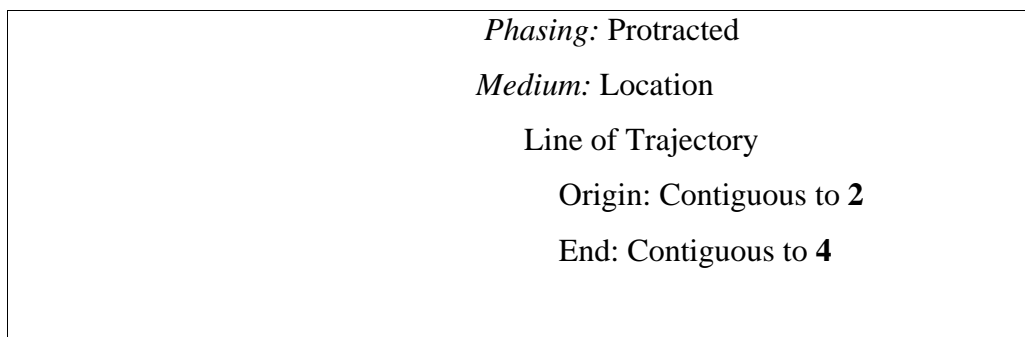


Fig.4 Basic cell of Verbs of Initial Contact and Ejection

As formally presented in Fig.4 above, the verbs from the present group lexicalize situations which can be regarded as consisting of two components (or sub-events), introduced as *Monodevelopment_a* and *Monodevelopment_b*, respectively. Thus, the main participants in the situation denoted by the verbs at hand are three. The first participant is marked with the set of values [*Source₁*, *Conditioner₁*, (*+Control*)₁] and its typical overt realization in active sentences is the subject position. As already discussed in section 4.1.2.1, a participant, which is specified for having a control over the event, can be also referred to as *Initiator*. Therefore, it is to be expected, that this participant (indexed with 1 in subscript in Fig.4 above) must also pertain to a set of semantic characteristics, usually restricted to [+Human(like)].

However, there are cases where the referent of the subject does not bear the [+Human(like)] feature. These are instances where the participant realized in the subject position is a complex mechanism, capable of running on its own, as illustrated in the examples in (61) below.

- (61) a. *Mašinata mjataše topkite za tennis (kato luda).*
 machine-the cast balls-the for tennis like crazy
 a' The machine was casting tennis balls (as crazy).
- b. *Mašinata xvūrljaše iskri.*
 machine-the threw sparks
 b' The machine was throwing sparks.

Nevertheless, a humanlike behaviour is implied, which is evident in the acceptability of the manner phrase *as crazy* in the example in (61a). Thus, these cases may be considered as examples of extending the basic meaning of the verbs at hand.

More such instances are illustrated in the examples in (63) below.

- (62) a. Vūlnite xvūrljaxa prūski ... (LCWB)
 waves-the threw splashes
 a' The waves threw splashes.
- b. ... kŕmata na sala se metna koso nagore ... (LCWB)
 stern-the of raft refl.cl. threw straight upward
 b' The stern of the raft rose straight up.

In all these cases, the referent of the subject may not have control over the situation lexicalized, but it is still marked by the features [Source₁, Conditioner₁].

Furthermore, the verbs *throw/xvūrljam* and *cast/mjatam* are also employed to lexicalize situations where no actual impact is implied as illustrated in the examples in (63) below.

- (63) a. ... the tenuous mist cast a sheen of silver. (BNC: A73 379)
 b. Kedrite xvūrljaxa ogormna sjanka. (LCWB)
 cedars-the threw an enormous shadow
 b' The cedars threw enormous shadow.

The lexicalization of similar situations with the use of extended verb meanings or metaphorical extensions of the verbs at hand is further discussed in section 4.3.

The second participant, marked with the subscript index 2, bears the value of *Source Extension* and it is generally not realized overtly with the verbs *throw* and *cast*, since it is only implicitly present, unless additionally specified as in the example in (64b) below.

- (64) a. ?? He threw the stones with hands.
 b. He threw the stones with his bare hands.

As already discussed in section 3.3.4, an element that is tightly incorporated in the meaning of a verb often cannot be expressed syntactically unless it provides us with additional information. This is illustrated in the different level of acceptability in the examples in (64a) and (64b) above.

The verb *shoot/streljam*, however, lexicalizes a situation where the participant that bears the value of *Source extension* is in addition specified for inherent properties, e.g. it must be a member of a semantically constrained set of objects including firearms and small arms. Hence, this participant is criterial and can be realized in the subject position (*Instrument Subject Promotion*), as illustrated in the examples in (65) below.

- (65) a. *A volley gun* cannot move and shoot in the same turn except to turn to face its target. (BNC: CN1 694)
 b. *Puškata* streljaše točno.
 rifle-the shot precisely
 b' The rifle was shooting precisely.

The third participant included in the lexical representation in Fig.4 above is marked with the set of values [*Absorber₃*, *Mover₃*] and it is syntactically realized as the direct object in active sentences or in the subject position of passive constructions as illustrated in the examples in (66) and (67) below.

- (66) a. Panicked, she threw *a chair* clumsily at him. (BNC: JYD 2366)
 b. He shot *a bullet* into the air.
- (67) a. *The bag of stones* was thrown into the air. (BNC: AMB 949)
 b. *The bullet* was shot at the back of the vehicle.

Besides, the beginning, the end, or the whole length of the *Trajectory Line* of the *Mover*₃ (also referred to as *Path* phrase or *Path* components) may sometimes be overtly realized in a prepositional phrase as illustrated in the examples in (68) below.

- (68) a. Another teenager was thrown *from the car*. (BNC: AJD 724)
 b. ... some old sandwiches, which I threw *into the water* for the fish.
 (BNC: ACK 2100)
 c. She threw him *over a wall*. (BNC: CH6 902)

In addition to the participants already discussed above, the verb *shoot*, rather uniquely, encodes also a participant that bears the value of *Limit* as demonstrated in the example in (69) below.

- (69) a. I should have let him shoot *the boy*. (BNC: H9N 235)
 b. *A giant* had just been shot. (BNC: F9Y 1260)

While English may use the verb *shoot* to lexicalize various aspects of the shooting event, Bulgarian employs different syntactic patterns and various aspectual prefixes⁶ to alter the meaning of the verb according to the situation it is taken to lexicalize, as illustrated in the examples in (70) below.

- (70) a. shoot a bullet – izstreljam kuršum (out-shoot bullet)
 b. shoot a gun – streljam s oružie (shoot with gun)
 c. shoot a man – prostreljam/zastreljam čovek (through-shoot/shoot dead man)
 d. shoot at the man – streljam po čoveka (shoot at man)

⁶ Cf. the discussions on the Bulgarian prefixes as morphological flagging in the mapping of semantic participants into syntax (Dimitrova-Vulchanova 1996/99) and on the semantics and the functions of the prefixes in Bulgarian (Guentcheva 2002).

The Bulgarian members of the *Verbs of Initial Contact and Ejection* group can also participate in the Absolutive *se*-construction (cf. the classification made in Dimitrova-Vulchanova 1996/99) where the *Source* is co-indexed with the *Absorber* of the *Force* and is overtly realized as the subject in the sentence. In this case, illustrated in the examples in (71) below, the verbs at hand again behave as *Manner of Motion* verbs.

- (71) a. Toj se xvŕli ot mosta.
 he refl.cl. threw from bridge-the
 a' He threw himself from the bridge.
- b. Toj se metna vŭv vodata.
 he refl.cl. cast into water-the
 b' He cast himself into the water.

This observation is in line with findings in Dimitrova-Vulchanova (1996/99) and recent experimental evidence in Dimitrova-Vulchanova et al. (in press) that force emission in self-conditioned events logically results in change of position/movement in space.

It is interesting to observe the differences between Bulgarian and English in the preferences of the overt realization of a *Path* component. In Bulgarian, phrases denoting *Path origin* (*from the bridge* in the example in (71a) above) and those denoting *Path end* (*into the water* in the example in (71b) above) are equally frequent. The preferred syntactic realization in the English corpora data, however, was primarily of a *Path end* phrase, as illustrated in the examples in (72a) and (72b) and rarely *Path length* phrase as in (72c).

- (72) a. Gerard threw himself *on the floor*. (BNC: K35 485)
 b. He rushed in and threw himself *into a chair*. (BNC: GW8 2384)
 c. Owen threw himself *over the wall* and dropped down. (BNC: HTX
 3395)

A necessary condition for this particular reading of the Bulgarian *se*-construction, namely, an Absolutive one, is the presence of an *Absorber*, since only the *Absorber*, and not the *Limit* in the situation is a potential *Monodeveloper*, hence it is *Criteria* and can be expressed as the syntactic subject in active sentences.

As the verb *shoot/streljam* can appear with both an *Absorber* and a *Limit*, its syntactic behaviour additionally proves this hypothesis. Only the *Absorber* of *shoot/streljam*, and not the *Limit*, can appear in this syntactic pattern, as reflected in the examples in (73) and (74) below.

- (73) a. Toj izstrelja topkata navisoko.
 he shot ball-the high up
 a' He shot the ball high up. (*the ball is Absorber*)
- b. Topkata se izstrelja navisoko.
 ball-the refl.cl. shot high up
 b' The ball shot high up.
- (74) a. Toj zastrelja čoveka.
 he shot man-the
 a' He shot the man. (*the man is Limit*)
- b. Čovekŭt se zastrelja. (Possible only on a reflexive reading)
 man-the refl.cl. shot
 b' The man shot himself.

It would be even more apparent if we use the same word (*čovek/man*) as the direct object in (74) and force it into a situation where it will correspond to the value of *Absorber*. For example, imagine the circus stunts where people are shot from cannon, then *the man* will behave exactly like *the ball*, both in real life, and syntactically. In fact, this very usage was found in the BNC, and is replicated in the example in (75) below.

- (75) ... nutters who want to be shot out of the mouth of a cannon ... (BNC:
 K4T 9712)

The change of verb meaning in accordance with the situation it denotes is further discussed in section 4.3 where I take into account some metaphorical situations that might be lexicalized by the verbs at hand. This approach is in line with Pustejovsky's view of a lexicon, based on the generativity of word senses, whereby the meaning of a word can be coerced to accommodate the sentential meaning (cf. section 2.3 for discussion and references), as well as with recent research on the effects of verb meaning on the verbs subcategorization preferences (cf. work by Hare et al. (2003), discussed in section 3.2).

4.2.3 *The Dual Lexicalization Pattern*

As mentioned earlier, some verbs can lexicalize either of the two basic event types discussed in sections 4.2.1 and 4.2.2, respectively. That is, these verbs can choose between the two basic lexical representations. This language irregularity is presently dubbed *The Dual Lexicalization Pattern* and is most probably based on the available conceptual representations of the events, which allow us to perceive the world from different perspectives, hence giving us, at least theoretically, an array of possible ways of parsing the events. Thus, the various patterns of lexicalization are accounted for in their correlation with concepts (e.g. situations).

As it was briefly discussed in section 4.2.2.2, a cross-linguistic variation is also observed. Thus, while a verb in a certain language is employed in the dual lexicalization pattern, other languages may use two separate items in denoting the two situation types, e.g. using separate lexicalizations.

Compare the situations lexicalized by the Bulgarian verbs *mušna* and *rŭgna* (both can be roughly translated as *poke, push, thrust, jab, or stick*) in the examples in (76) and (77) below.

- (76) a. Toj mušna/rŭgna mŭža v gŭrdite.
 he poked man-the in chest-the
 a' He poked the man in the chest.

- b. Toj se mušna/rŭgna [s prŭst] v gŭrdite.
 he refl.cl. poked [with finger] in chest-the
 b' He poked himself in the chest. (found also in BNC: CR6 3264)
- (77) a. Toj mušna/?rŭgna⁷ rŭka v džoba j.
 he slipped/stuck hand in pocket-the poss.cl.fem.
 a' He slipped/stuck a hand into her pocket.
- b. Rŭkata mu se mušna/?rŭgna v džoba j.
 hand his refl.cl. slipped in pocket poss.cl.fem.
 b' His hand slipped into her pocket.
- c. Toj se mušna/rŭgna v tŭlpata i izčezna.
 he refl.cl. slipped in crowd-the and disappeared
 c' He slipped into the crowd and disappeared.

In the examples in (76), both *mušna* and *rŭgna* behave as *Verbs of Contact* (very similar to *stab/proboda*). The direct object in (76a) carries the value of *Limit* and when the verbs are used in a *se*-construction (as in (76b)), the sentence acquires a reflexive meaning.

On the other hand, considering the syntactic patterns displayed in the sentences in (77), the same verbs seem to side together with the verbs denoting a conditioned event. The direct object in the example in (77a) bears the value of *Absorber* and it can appear as the subject of an Absolutive *se*-construction as illustrated in the example in (77b). In addition, the example in (77c) demonstrates a syntactic pattern parallel to the one displayed by *Verbs of Initial Contact and Ejection* (illustrated in the examples in (71) above.)

⁷ Compared to *mušna*, *rŭgna* is not so frequent, rather obsolete, and with a slightly negative connotation, which does not always result in ungrammaticality but makes the usage of *rŭgna* quite inappropriate in many sentences.

However, these verbs are not unique. Two other Bulgarian verbs – *blūsna* (*knock down*) and *butna* (*push*) – display similar syntactic behaviour due to their ability to lexicalize events of both types. On the one hand, they can denote a *Contact* situation where a force is released by the *Source* and is absorbed by a *Limit* through its contact with the *Source* or a potential *Source extension*. The event is accomplished when contact is attained. In this case, the lexical representation of these two verbs (their *cell*) will be similar to the one already observed for the verbs from the first group (Fig.1) and they will display analogous syntactic behaviour. On the other hand, the verbs can also denote a *Conditioned* event where the element absorbing the force (here an *Absorber*) undergoes a *Monodevelopment* with *Medium:Location* as a result of the energy received through the contact. In this case, their lexical representation will resemble the abstract cell in Fig.2, but will differ in *Aspectuality* [+2-point] and *Consistency* [+contact]. Whether these verbs lexicalize an event of the first type or the second may be checked by introducing a continuation, implying no consequent movement, which would be possible only with the first type of situation, but not with the second, as illustrated in the examples in (78a) and (78b), respectively.

- (78) a. Toj ja blūsna/butna, no tja ne pomrūdna.
 he her knocked/pushed, but she not moved
 a' He pushed her but she did not move.
- b. Toj ja blūsna/butna na zemjata, *no tja ne pomrūdna.
 he her knocked/pushed on ground, but she not moved
 b' He knocked her down, *but she did not move.

In addition, some verbs, which primarily lexicalize one type of event, may under certain conditions lexicalize an event of the other type. Let us take for example the verbs from the first group, *Verbs of Contact*.

In one particular case, the Bulgarian verb *udarja* (*hit*) lexicalizes an event where the *Source* is recognized as an indivisible entity and is criterial only as such, e.g. the

Source acts as if there is no *Source extension* at all⁸. Therefore, no part of the *Source* may be overtly specified. Since contact must be attained, as it is an essential part of the event lexicalized, and there is no separate participant marked as *Source extension*, we must conclude that it is the *Source* itself, which is set into motion in order to attain contact. Thus, we have only one participant *i* that combines the three values [*Conditioner_i*, *Source_i*, *Mover_i*]. This is an example of a self-conditioned event as illustrated in the sentences in (79) below.

- (79) a. Toj se udari v stenata [*s krak/rūka/glava].
 he refl.cl. hit in wall-the [with foot/ hand/head]
 a' He bumped into the wall [* with his foot/hand/head].
- b. Toj udari stenata [s krak/s rūka/s glava].
 he hit wall-the [with foot/ hand/head]
 b' He hit the wall [with his foot/hand/head].

The examples (79a) and (79b) differ in that (79b) allows for a *Source Extension* completion, while the sentence in (79a) is ungrammatical if such a phrase is added to it. However, a continuation, which does not refer to part of the *Source*, may be added as illustrated in the example in (80) below.

- (80) a. Toj se udari v stenata s novata si kola.
 he refl. hit in wall-the with new-the refl.poss.cl. car
 a' He bumped into the wall with his new car.

This is possible because this type of continuation does not directly conflict with the *Indivisibility Constraint* we have on the *Source*. Therefore, the *Source* is still criterial by *Contact*, and the continuation should be essentially regarded as denoting

⁸ Similar behaviour is observed with verbs like *jump* and *leap* and is called the *ultimate shrinking of a chain* (cf. Dimitrova-Vulchanova 1996/99).

Manner, rather than an *Instrument/Source extension*, as illustrated in the examples in (81a) and (81b) below.

- (81) a. Toj se udari v stenata [s glavata napred].
 he refl. cl. hit in wall-the [with head-the ahead]
 a' He bumped into the wall his head ahead.
- b. Toj se udari v stenata [s vsička sila].
 he refl. cl. hit in wall-the [with all strength]
 b' He violently bumped into the wall.

Thus, the abstract cell of *udarja* (*hit*), when lexicalizing a self-conditioned event, must be given the formal representation illustrated in Fig.5 below.

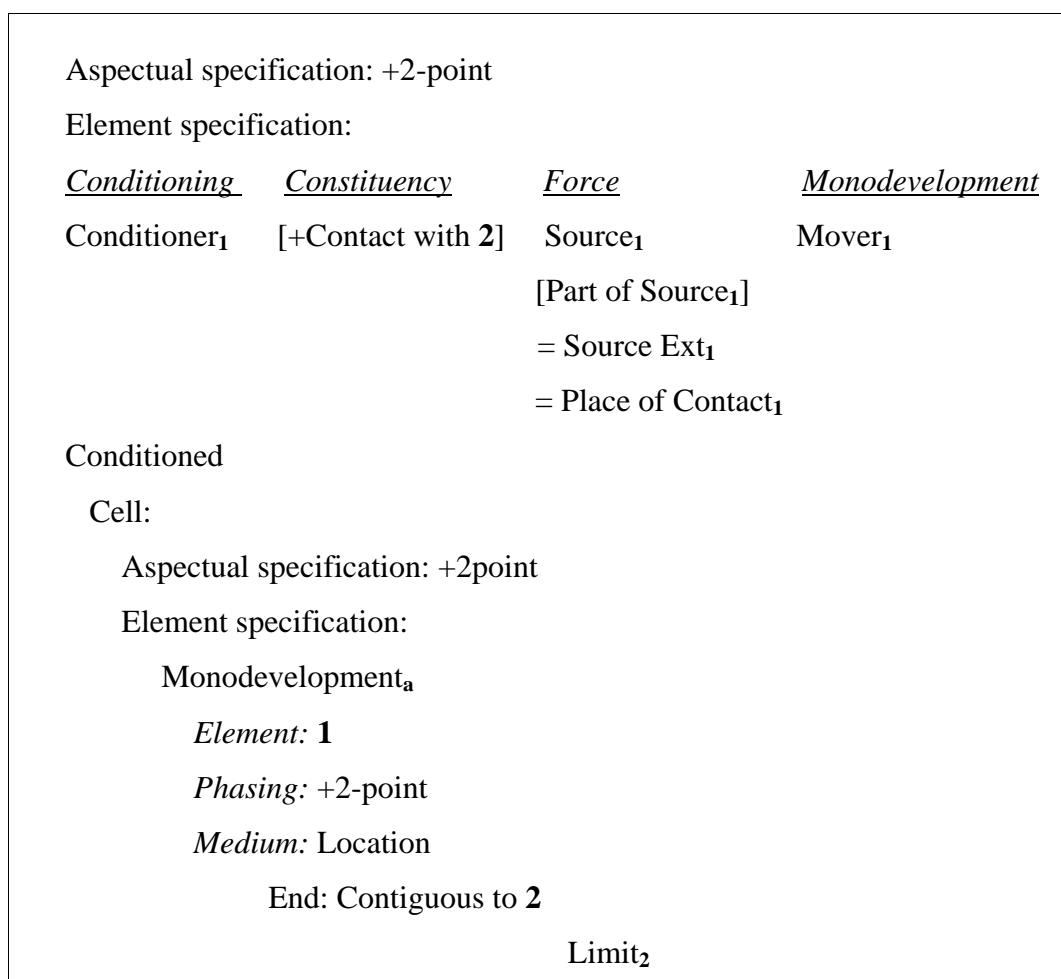


Fig.5 Basic *cell* of *udarja* (*hit*) lexicalizing a self-conditioned event

Thus, the only way to have syntactically realized part of the *Source* is by means of the *Place of Contact Promotion* discussed in section 4.2.1. Hence, a part of the *Source*, instead of the *Source* itself, may emerge in a subject position. Yet, it is not *criterial* unless the situation is exhausted in reaching a *Contact*, as we can see in the examples (82a) and (82b) below.

- (82) a. *Rŭkata/glavata mu se udari.
 hand-the/head-the his refl.cl. hit
 a' *His hand/head hit itself.
- b. Rŭkata/glavata mu se udari v stenata.
 hand-the/head-the his refl.cl. hit in wall-the
 b' His hand/head hit into the wall.

As already discussed above (section 4.2.1), most of the Bulgarian members from the *Verbs of Contact* group differ from *udarja/hit* with respect to the permitted reading of *se*-construction. However, they can also lexicalize an event from the opposite type (e.g. a conditioned event) if used as intransitives. This is a crucial point where English and Bulgarian differ typologically. While in English the process of transitivization is widespread as for example *walk* in *He walked the dog*, in Bulgarian the reverse practice is more common. Verbs that are essentially transitive are used as intransitive to denote an event of motion performed in a certain manner, as illustrated in the examples in (83) below.

- (83) a. Knigata pljasna *[na zemjata].
 book-the slapped on ground-the
 a' The book slapped *[on the ground].
- b. Topkata capna *[vŭv vodata].
 ball-the smashed in water
 b' The ball fell *[in the water].

Thus, these *Verbs of Contact* change into *Manner of Motion* verbs, hence they not only allow for an *End of Path* element but also require it, as it is now *Criteria* by contact, which is initially encoded in their meaning. This is demonstrated in the unacceptability of the sentences in (83) without the prepositional phrases specifying the *Contact* point. In Fig.5 above, this is the element marked as *Limit₂*, and the end of the trajectory line (*End of Path*) is marked as *Contiguous to 2*.

The syntactic behaviour displayed by all the *Verbs of Force Emission* discussed in this work, as well as the types of situations they can lexicalize, suggest that we should consider these verbs to belong conceptually to a larger group of *Verbs of Motion*.

4.3 Extended uses of the examined verbs

Up to this point, the analyses presented in section 4.2 focused on verbs used only in the sense of *Concrete Physical Impact*. The present section explores some of the possible extended uses of the investigated verbs as encountered in the corpora. The treatment of verbs in sentences with *Extended Meaning* builds on the compositional interpretation of words in context based on their QUALIA STRUCTURE and the generative semantic operations (more specifically TYPE COERCION and CO-COMPOSITION), as proposed by Pustejovsky (1995) and discussed in section 2.3 of the present work.

The analyses of the available corpus data showed various degrees of usage of the different verbs in sentences with *Extended Meaning*, ranging from 0% for the English verb *dab* and the Bulgarian verbs *potupa* (*tap*) and *mušna* (*thrust*) to 67% for the English *cast* and 70% for the Bulgarian verb *teglja* (*drag*). It is worth mentioning that although not entirely equivalent, the English and the Bulgarian verbs displayed strong similarities in the situations a particular pair of correlates could lexicalize cross-linguistically.

A sentence was counted as displaying *Extended Meaning* of the head verb if one or more of the syntactically realized participants did not conform to the initial lexical representation of the verb at hand, thus, causing a deviation from the basic meaning of the verb as denoting a situation of *Concrete Physical Impact*. This is demonstrated in the English examples in (84) and the Bulgarian examples in (85), where the non-typical participants are given in italics.

- (84) a. She cast such *thoughts* from her. (BNC: JXT 1115)
 b. Isabel shot her *a sharp glance* over the cards ... (BNC: AD1 1610)
 c. *A pang of yearning* stabbed her ... (BNC: H7W 1078)
 d. *A headache* would drag on for days. (BNC: CB0 703)
 e. ... he was suddenly struck by *inspiration*. (BNC: A7H 681)
- (85) a. Tja hvūrli *begūl pogled* na platnoto vūrXu trinožnika ... (LCWB)
 she threw cursory glance at canvas-the on tripod-the
 a' She threw a passing glance at the canvas on the tripod.
- b. *Jarkite oči* se metnaxa naljavo, nadjasno ... (LCWB)
 bright-the eyes refl.cl. cast left right
 b' The bright eyes were shooting left and right ...
- c. ... i v gūrdite go probode *ostra bolka*. (LCWB)
 and in chest-the him stabbed sharp pain
 c' And a sharp pain stabbed him in the chest.
- d. Drugi vlačat *duševnata si mūka* s desetiletija. (LCWB)
 other drag mental refl.poss.cl. suffering with decades
 d' Others are dragging their mental sufferings for decades.
- e. ... *momentūt* beše ucelen dobre. (LCWB)
 moment-the was struck well
 e' The precise moment was hit.

The examples in (84) and (85) demonstrate instances of interface between the semantic characteristics of the verbs at hand and the NPs. Thus, they are to be considered as metaphorical extensions of the meaning of the verbs and we do not need to propose additional senses to the basic meaning of the verb at hand. Instead, these can be treated by a semantic transformation called TYPE COERCION, which, as already discussed in section 2.3, is *a semantic operation that converts an argument to the type*

expected by a function, where it would otherwise result as a type error (Pustejovsky (1995), p. 111.) Thus, although the qualia structures of the non-canonical participants the examples in (84) and (85) above would not conform to the type expected by the verb, an application of the TYPE COERCION operation accounts for the acceptability of these sentences.

In addition, in cases of metaphorical extensions, the verbs generally pertain to their common syntactic patterns as demonstrated in the examples in (84e) and (85e), displaying passive constructions in English and Bulgarian, respectively, and in the Bulgarian examples in (86) below, illustrating the *Instrument Subject Promotion*.

- (86) a. ... *s moite oči šte go streljam* ... (LCWB)
with my eyes will him shoot
a' I will shoot him with my eyes.
- b. ... *krasivite oči na momčeto streljaja zloba i nenavist*. (LCWB)
beautiful-the eyes of boy-the shot spite and hatred
b' The boy's beautiful eyes were shooting ill will and hatred.

The same diatheses were found with the *Concrete Physical Impact* sense of the verb *streljam/shoot*, as already mentioned in section 4.2.2.3 above, and illustrated in the corresponding examples in (87) below.

- (87) a. ... *streljaja v tulpata s puškite si*. (LCWB)
shot in crowd-the with rifles-the refl.poss.cl.
a' They were shooting at the crowd with their rifles.
- b. ... *puškata streljaše patroni 22-ri kalibūr*.
rifle-the shot bullets 22nd gauge
b' The rifle was shooting bullets gauge 22.

However, there are cases where the subtle differences in the situation denoted by the verb at hand are related to rather distinct lexicalization patterns and differences in

the semantic features of the participants involved. Such a case was mentioned previously (in section 4.2.2.3) with the verbs *throw/xvŭrljam* and *cast/mjatam* when lexicalizing an event where no actual impact is implied, as illustrated in the examples in (88) below.

- (88) a. *Sgradata xvŭrljaše sjanka vŭrxu blizkata gradina.*
 building-the threw shadow over near-the garden
 a' The building was throwing shadow over the neighbouring garden.
- b. *The street light* threw strange shadows among the hoardings. (BNC: CE9 692)

In the situations lexicalized in the examples in (88), the *Ejection* component, typically present in the semantic representation of the verb *throw*, must be substituted by an *Emission/Projection* component, since no real *throwing*, in the *Concrete Physical Impact* sense of the verb, occurs. Thus, any element capable of producing a shadow either by emitting light on to an object/obstacle (as demonstrated in the example in (88b)) or by being such an object/obstacle itself (as illustrated in (88a)) can be realized in the subject position of a sentence denoting this type of situation. Therefore, no participant marked for [+Control] must be included in the lexical representation of *throw* in this case.

Instead of positing a completely new meaning of the verbs *throw/xvŭrljam* and *cast/mjatam* when lexicalizing an event of emitting/projecting a light/shadow, the treatment adopted here involves the semantic operation called CO-COMPOSITION in a way similar to the treatment presented for the verb *bake* (section 2.3.2.3), however slightly more complex.

The process of CO-COMPOSITION involves shifting of the event type of the head verb as a result of the information carried by its complements which act on the verb and take it as an argument (Pustejovsky, 1995). In this particular case (the examples in (88) above), the FORMAL quale of the participant realized in the subject position does not bear the [+Human(like)] feature, as expected by the lexical representation of the verb, discussed in section 4.2.2.3 and is of different event type. Therefore, a semantic

operation must take place, allowing for semantic interpretation from both the head verb and its complements.

Cross-linguistically, the two languages employ the same lexical items to lexicalize this situation. It must be mentioned, however, that the Bulgarian verb *mjatam* (*cast*), as opposed to *xvŕljam* (*throw*), is also marked for [+Force] in its *Concrete Physical Impact* sense, thus making it difficult, if not impossible, to be used in lexicalizing an effortless situation such as *throwing a shadow*, as illustrated by the unacceptability of the example in (89) below.

- (89) **Sgradata* mjataše sjanka.
 building-the cast shadow
 The building was casting a shadow.

Yet, if used in a context, where some achievement is implied, thus entailing use of force (even though only metaphorically), *mjatam* (*cast*) can also lexicalize the same situation, as illustrated in the example in (90) below.

- (90) *Dŕrvoto* mjataše sjankata si *čak do drugija brjag na rekata*.
 tree-the cast shadow-the refl.poss.cl. as far as to other-the side of river-the
 The tree cast its shadow *all the way to the other side of the river*.

The examples in (89) and (90) above demonstrate again a positive evidence for the interface between the semantic characteristics of the verb and its complements⁹ and interpretation from both.

⁹ As already discussed in section 3.1, the lexically encoded semantic participants in the situation denoted by a verb, are considered its complements. Since the representation of *mjatam* (*cast*) includes a *Path* component, we can consider *the End of Path* phrase in (90) as a complement which may alter the basic meaning of a verb (cf. Pustejovsky (1995) for a similar treatment of *float in the cave*).

4.4 Conclusions

This chapter presented the results from corpus data analyses of a set of English verbs and their correlates in Bulgarian. The investigation aimed at exploring the syntactic distribution of a group of verbs in relation to their semantic properties.

Based on the analyses of the English and the Bulgarian corpus data, I have outlined the basic lexicalization patterns of the examined verbs and grouped them according to the types of situations they could denote. Hence the verbs were not simply enumerated in various lists but distributed in a more net-like pattern, accounting for the possibility of one verb to lexicalize situations of different types. Thus, the lexical representations of the verbs at hand were discussed in relation to the interface of conceptual structure with syntax.

Special attention was paid to the event segmentation and how it may correspond to lexical items cross-linguistically. It was observed that languages may vary in the mapping of event components into lexical items.

In addition, some metaphorical uses of the examined verbs were discussed in relation with the possible extensions of the verb basic meaning as result of the interaction of the semantic features of the verb and the information carried by its complements.

5. Online Sentence Continuation Studies

Experimental evidence in recent research by Koenig et al. (2002, 2003), discussed in Chapter 3.1, provides positive indication that the information which is lexically encoded in a verb is accessed immediately upon recognition of that verb, resulting from reading or hearing it. Therefore, it is more likely for lexically encoded participant information to occur in the continuation of a sentence headed by the verb.

To test native speakers' intuition about the most prominent participants in the various situations denoted by the set of target verbs, two online sentence continuation studies have been conducted as part of my research project. As the name suggests, the experiments were designed as a set of unfinished sentences, which the subjects had to complete, if they considered it appropriate. Thus, the participants¹ in the tests were neither prompted to give a certain type of continuations, nor pressed to provide a continuation by all means.

The main idea behind these experiments was to test whether the participant information believed to be lexically encoded in the conceptual representation of the target verbs from the corpora studies do play role in the online processing of sentences. Thus, I expected to receive a significantly higher percentage of completions related to semantic *participant* information which is lexically encoded, than the percentage of the responses that do not include lexically encoded *participant* information.

¹ To avoid terminological confusion in this section, I will refer to the participants in the tests as subjects or participants, while I will italicize *participant*, when referring to the information encoded in verbs.

5.1 Design and Methodology

The studies comprise two parallel continuation tests, aimed at assessing the psychological reality (as described in Koenig et al. 2002 and discussed earlier in section 3.2) behind the formal lexical representations suggested in the previous chapter for the target group of corresponding verbs in the two languages – English and Bulgarian. For each of the analysed verbs, a preliminary research was conducted, using both corpus data and the results of pilot continuation tests.² Based on this preliminary research, I have outlined the basic representation for each verb, and grouped the verbs together according to their syntactic behaviour and the situations they can lexicalize, as discussed in Chapter 4. Thus, the most prominent syntactic patterns displayed by the verbs at hand were used in designing the sentences to be completed.

The methodology used in these studies generally followed the one used by Koenig and his colleagues in their research discussed in section 3.2. When subjects are asked to complete sentences, the expectations are that *participant* information which is lexically encoded in the meaning of the verb is more likely to be expressed overtly, since it is retrieved upon recognition of that verb. Thus, we assume that if a *participant* is part of the lexical representation of a verb, information about it is already activated while parsing the verb at hand and it becomes more prominent and likely to be used in continuing the sentence headed by the verb.

5.1.1 Participants

The participants in the tests were all native speakers of English or Bulgarian, respectively. They ranged from twenty to fifty years of age and they had different background with regard to place of origin and education. Each of the participants completed the test on his/her own with no additional instructions besides the ones included in the test. Thus, they were asked only to read the instructions on the screen and then proceed with the task.

² The pilot studies were carried out on paper and included the results of twelve participants for English and twenty-four for Bulgarian.

Each of the tests was presented to thirty-five to forty subjects. However, some of the participants did not complete the test in its whole (e.g. they have stopped before the end of the test) and their results were discarded. For the final analyses I have used the results of thirty subjects for each of the tests.

5.1.2 *Stimuli*

There were two parallel sets of stimuli for English and Bulgarian, respectively. Each set consisted of fifty sentences, containing as many as eighteen target verbs, together with the same amount of filler sentences containing various distracter verbs. Thirty³ of the target sentences consisted only of a subject and a verb, while the rest of the target sentences contained also a direct object. The filler sentences were designed in the same way, so that no apparent distinction between target and distracter verbs could be made throughout the test.

Eighteen of the sentences, consisting of a subject and a verb, presented a basic situation lexicalized by the verb, where the subject carried the value of *Initiator*, as illustrated in the English examples in (1a, b) and the respective Bulgarian examples in (1a', b') below.

- (1) a. Sam kicked ____
 a' Stefan ritna ____
 b. Lucy slapped ____
 b' Lili pljasna ____

Another twelve sentences, syntactically constructed also as subject plus verb only, differed in that they presented non-canonical situations lexicalized by the verbs at hand, where the subject could not be assigned the value of *Initiator*, but carried instead

³ The number of the sentences may differ by one or two between the English and the Bulgarian tests. This is due to a common characteristic of Bulgarian to display a two-verb split, e.g. for one English verb there might be two Bulgarian verbs with almost overlapping definitions (cf. Dimitrova-Vulchanova et al. (in press) for establishing the correlation with verbs of biological motion).

the value of *Source*, *Source Extension*, or *Absorber*, as illustrated in the examples in (2) to (4), respectively.

- (2) a. Vŭlnite pljaskaxa ____ (Source)
 waves-the slapped
 a' The sea was slapping ____
- (3) a. Nožŭt režeŭe ____ (Source Extension)
 knife-the cut
 a' The knife cut ____
- (4) a. Tazi povŭrxnost se draska ____ (Absorber)
 this surface refl.cl. scratches
 a' This surface scratches ____

The examples in (2a) to (4a) present the sentences included in the Bulgarian test, while the examples in (2a') to (4a') are the sentences included in the English test.

The last twenty sentences were similar to the first ones in that their subject carried the value of *Initiator*, but in addition these sentences had also a direct object.⁴ Thus, they were virtually completed, e.g. they were all grammatical without any obligatory completion needed. Illustrative examples from Bulgarian (5a, b) and English (5a', b') are given below.

- (5) a. Toj otrjaza dŭrvoto ____
 he perf.-cut tree-the
 a' He cut the tree ____
- b. Stojan udari vratata ____
 Stojan hit door-the
 b' Mark hit the door ____

⁴ Two stimuli (one in English and one in Bulgarian) had a prepositional phrase instead.

As previously mentioned above, all the filler sentences were constructed in a similar way, so that they structurally resembled the target sentences and were equal in number. Sample sentences of the ones used as fillers are presented in (6) below.

- (6) a. Toj ču ____ *Subject[Initiator] Verb* ____
 a' He heard ____
- b. Vjatūrūt otnese ____ *Subject[Source] Verb* ____
 wind-the perf.-carried
 b' The wind carried ____
- c. Valja izpi kafeto ____ *Subject[Initiator] Verb Object* ____
 Valja drank coffee-the
 c' I drank my coffee ____

As illustrated by the examples given so far, most of the sentences included in the tests were formed with verbs in The Simple Past Tense. However, for some English verbs, sentences in The Past Progressive were also included to test whether there would be a significant difference in preferred continuations. In Bulgarian, this was achieved by including different aspectual forms of the respective verbs as illustrated in the examples in (7) below.

- (7) a. Bob hit ____
 a' Toni udari ____
 b. Peter was hitting ____
 b' Petūr udrjaše ____

Since this was not the main objective, not all of the possible aspectual variances were included. These sample sentences were intended only as a preliminary attempt to elicit experimental evidence on how aspect can modify meaning, hence the displayed syntactic patterns (cf. Dimitrova-Vulchanova 1996/99 for theoretical assumptions on

verb aspectuality and diathesis). Thus, my primary aim was to check for the more basic situations and the information encoded in verbs denoting them.

The only English example in The Simple Present Tense was the one presented in (4a) above as it involves a particular syntactic construction known as *The Middle Alternation* (Levin 1995). The sentence denotes a situation where the *Absorber* (realized as the subject of the sentence) is affected when no apparent *Source* or *Initiator* seems to be involved. Its semantic equivalent in Bulgarian is syntactically realized with an Absolutive *se*-construction as illustrated in (4a') above.

In addition, there is not a complete overlap between the verbs included in the corpora analyses and those discussed in the tests results. This is due to various reasons as for example some of the verbs were added in a later stage of the research process or the target sentences elicited more continuations with the verb's homonym like the case with the Bulgarian verb *smazvam/smaža*, which was intended in the sense of *smash*, but it received quite as many continuations with the meaning of *oil/grease* (which was not unexpected, though).

5.1.3 Procedure

The participants in the tests were asked to “complete the sentences where appropriate, without spending too much time on any of the items.” Thus, the subjects were encouraged to write down each continuation fast; so that it would be the first “thing” that came into their mind (additional literature on the methodology of similar type of tests can be found in Koenig et al. 2002, 2003).

Once they had read the instructions, the participants were provided with only one sentence at a time, which they had to read and decide whether to continue or not. After the decision (and an eventual completion) they had to press a "send" button in order for the next sentence to appear, while the previous one disappeared. This was designed so that a minimum interaction between the information activated from the different sentences could be achieved.

The order of the sentences presented to the subjects was different each time the test was started. A randomly generated sequence of all the sentences – both stimuli and

fillers, was created in the beginning of every trial. Thus, we aimed at eliminating any eventual side effects which could be caused by a strict order of the sentences. In addition, the participants could not go back and see their results until they have been through all the sentences in the test, thus completing it in its whole. However, no further corrections could be made.

5.2 Analyses of the Results

The results from the continuation tests were assessed identically to the corpora data analyses discussed in Chapter 4, however, with a focus on the semantics of the continuations provided. Thus, again, I distinguished between sentences denoting a situation of *Concrete Physical Impact* and those implying an *Extended Meaning* of the verb at hand.

As for the syntactic approach, I singled out only the direct object (when it was part of the continuation given by the subjects) while the rest of the continuations were collapsed in a single column, named *Complementation*, regardless of whether they were overtly expressed as clitics, prepositional phrases, or adverbials. Thus, I concentrated on the semantic features of the *participants* elicited in the responses. Therefore, it was of no relevance whether exactly the same lexical items were used as long as they displayed the same semantic features. The methodology of the analyses is illustrated in the examples⁵ in (8) and (9) below.

- (8) a. Bojan probode *s nož pūržolata*.
 Bojan stabbed with knife steak-the
 a' Bojan stabbed *the steak with a knife*.
- b. Bojan probode *balona s igla*.
 Bojan stabbed balloon-the with needle
 b' Bojan stabbed *the balloon with a needle*.

⁵ All the examples in this chapter are taken from the results of the tests for English and Bulgarian, respectively.

- (9) a. Jordan blūsna *Marija po glavata*.
 Jordan knocked Maria on head-the
 a' Jordan knocked Maria on her head.
- b. Jordan blūsna *količkata s xranata kŭm Vanja*.
 Jordan knocked cart-the with food towards Vanja
 b' Jordan pushed *the shopping cart towards Vanja*.

The continuation in (8a) is syntactically realized as [PP O_d], while the one in (8b) can be formally presented as [O_d PP] and all these positions are filled by different lexical items. However, the direct objects in both sentences carry the value of *Limit*, following the formal representation of the verb *proboda* (*stab*), discussed in section 4.2.1. Similarly, the prepositional phrases in the two continuations are assigned the value of *Source Extension*.

In the examples in (9), we again have a combination of a direct object plus prepositional phrase. However, the referents of these constituents differ semantically. While the direct object in (9a) bears the value of *Limit*, the direct object in (9b) is assigned the value of *Absorber*. In addition, the prepositional phrase in (9a) is characterized as *Place of Contact*, while the prepositional place in (9b) is identified as a *Path* component (more specifically *Orientation*).

The analyses of the results from the two experiments are summarized in separate tables for each language, presented in Appendices D and E, for English and Bulgarian, respectively. The figures in these tables are given in numbers and not in percentage.

In addition, it is worth mentioning that although the subjects were not asked to complete every sentence by all means, there were 25 empty fields⁶ for Bulgarian (which is approximately 2% of the answers taken into account for Bulgarian) and only 8 for English (which is slightly more than half percent of the analyzed answers for English).

Moreover, the continuations which consisted only of conjunctions introducing another sentence, were singled out in a separated column, named *Conjunction*, since they were not considered as directly related to the semantics of any of the *participants*

⁶ In the tables in the Appendices these are given in a separate column, named *No Continuation*.

in the situation denoted by the verb at hand. However, there were only 32 such continuations for Bulgarian (2,5%) and 38 for English (2,8%). Therefore, we must conclude that most of the continuations elicited in the responses of the subjects, were highly influenced by the information that was activated upon recognition of the verbs.

As the objective of the online continuation studies was to check for potential semantic participants, encoded in the lexical representation of the verbs at hand, the discussion of the results concentrates on the assessment of the semantic features displayed by the participants in the continuations provided.

5.3 Results and Discussion

As already discussed, a basic formal representation was established for each group of verbs examined (illustrated in figures 1 to 5 in Chapter 4). In addition, a more detailed representation was created for each of the verbs, including specific features (if any) of the verb at hand. These representations were based on the lexicalization patterns displayed by the verbs in the corpus data analyzed and the semantic features ascribed to the participants in the situations as lexicalized by the verbs at hand. The online continuation studies aimed to provide experimental evidence for or against the adopted format of lexical representation and the information that it encoded.

The overall results⁷ indicated, indeed, that most of the elicited continuations referred to lexically encoded semantic participants. For a more comprehensible overview of the results, the discussion follows the grouping of the verbs presented in Chapter 4 and it takes into account the three different types of stimuli, as presented in section 5.1.2 above and in the Tables of Results in Appendices D and E, for English and Bulgarian, respectively.

⁷ The percentages presented in the discussion are calculated for groups of verbs. However, the results are even more significant if each verb is taken on its own. The numbers for each verb are given in the relevant tables at the end of this work.

5.3.1 Results for Verbs of Contact

The first group of verbs to be discussed are *Verbs of Contact*. These are the English verbs *hit*, *strike*, *tap*, *slap*, *stab*, and *kick* and their Bulgarian correlates *udarja*, *ucelja*, *potupam*, *pljasna*, *proboda*, and *ritna*. In addition, the verbs *knock* (*tropam*) and *smash* (*capna*), as well as the Bulgarian verb *blūsna* (*push*), are also included in the discussion when lexicalizing an event where only contact is implicated (e.g. if the participant realized in direct object position is marked with the value of *Limit*).

5.3.1.1 Results for stimuli of the type *Subject*_[Initiator] *Verb* ____

Following the lexical representation of these verbs and having provided the participant with the value *Initiator* (cf. the stimuli in Table 1 in Appendices D and E for English and Bulgarian, respectively), the expectations were mostly for continuations with a participant marked as *Limit* realized either in the direct object position or in PP_{on/at}. In addition, a relatively high percentage of continuations that could be assigned the values of *Source Extension* or *Place of Contact* was also expected.

As it can be seen in the results displayed in Table 1 of Appendices D and E (for English and Bulgarian, respectively,) the prevailing continuations received for the first group of verbs for the stimuli of the type *Subject*_[Initiator] *Verb* ____ related to participants with the value of *Limit*.

Thus, 57% of the sentences⁸ (both in English and in Bulgarian) elicited a continuation counted as *Limit* and realized in direct object position, as illustrated in the example in (10) below.

⁸ As there were multiple continuations (containing more than one phrase/participant), the percentages were not calculated against the total number of phrases received and analyzed. Instead, a more truthful approach was to count the sentences which elicited a participant with a certain value and calculate the percentage against all the sentences within the group discussed. Thus, one sentence could be included in the calculation of the results for two different values if it had been continued with two phrases, as illustrated in i) below.

- i. Bob hit the ball with the bat.

Therefore, the total sum in percentage would not necessarily equal one hundred.

- (10) a. Bob hit *his brother*.
- b. Toni udari *sestra si*.
Toni hit sister refl.poss.cl.
- b' Toni_i hit his_i sister.

This was to be expected since the stimuli contained transitive verbs and most of them were not grammatical unless provided with a direct object.

However, there were another 23% of the English and 19% of the Bulgarian sentences that received continuations regarded as *Limit*, but syntactically realized with a prepositional phrase, as exemplified in the sentences in (11) below.

- (11) a. Bill stabbed *at the steak*.
- b. Toni udari *po masata*.
Toni hit on table-the
- b' Toni hit the table.

The majority of these continuations were given for the stimuli denoting an iterative process (*Iteration*), as illustrated in the examples in (12) and (13) for English and Bulgarian, respectively.

- (12) a. The girl was knocking *on the hollow tree*.
- b. Tom tapped *on the window*.
- (13) a. Petūr udrjaše *po stenata*.
Peter hitting on wall-the
- a' Peter was hitting the wall.
- b. Momičeto tropaše *po masata*.
girl-the knocking on table-the
- b' The girl was knocking on the table.

These results confirmed the expectations with *Verbs of Contact* to elicit continuations related to participants with the value of *Limit* as predicted by the model (cf. the discussion in section 4.2.1). However, the presence of *Iteration* may substitute for the *Criteriality* of the second participant in *Contact* relation (e.g. the *Limit*),⁹ which now can become implicit and must not be syntactically realized as illustrated in the examples in (14) below.

- (14) a. The girl was knocking vigorously.
 b. Momičeto tropaše nervno.
 girl-the knocking nervously
 b' The girl was knocking nervously.

Yet, more empirical studies designed to assess these subtle differences in verb meaning are necessary to outline the details in the representations of verbs as denoting such situations (e.g. including *Iteration*) within the language, as well as cross-linguistically.

Continuations specified for *Place of Contact* were given to 13% of the Bulgarian sentences and 10% of the English ones, from which less than 2% (for both tests) were realized as direct object. These are illustrated in the examples in (15) and (16) below (for English and Bulgarian, respectively).

- (15) a. Bob hit himself *over the head*.
 b. Bob hit *his head*.
- (16) a. Filip uceli točno *centŭra na mišenata*.
 Philip struck straight centre-the of target
 a' Philip struck right into the centre of the target.

⁹ Cf. Hellan & Dimitrova-Vulchanova (2000) for analysis and discussion on *Criteriality* and grammatical realization of *Criterial* participants.

- b. Filip uceli mišenata *v sredata*.
Philip struck the target in middle-the
b' Philip struck the target in the middle.

As expected, a relatively high percentage of the sentences elicited continuations with a participant marked as *Source Extension* – 28% for Bulgarian and 14% for English, demonstrated in the examples in (17) below.

- (17) a. Bob hit Mary *with a baseball bat*.
b. Toni udari silno *s prűčkata*.
Toni hit hard with stick-the
b' Toni hit hard with the stick.

Bulgarian differed considerably from English in that there was no participant with the value of *Source Extension* realized in the direct object position, while in English, in more than one third of the sentences with *Source Extension*, this participant was realized as the direct object, like illustrated in the examples in (18) below.

- (18) a. Tom tapped *his finger*.
b. Bill stabbed *his pencil* into the rubber.

This find is very interesting regarding the fact that Bulgarian also allows for syntactic realization of the participant with the value of *Source Extension* in the direct object position, as illustrated in the examples in (19) below.

- (19) a. Stiga si rital *tezi kraka*.¹⁰
enough refl.cl. kick these feet
a' Stop kicking your feet.

¹⁰ Cf. Dimitrova-Vulchanova (1996/99, 2001) for discussion on the possible syntactic patterns predicted by the model for the verb *kick* in English and Bulgarian.

- b. Toj udari *rŭka/jumruk* v masata.
he hit hand/fist in table-the
- b' He his hand/fist onto the table.

However, the tendency in Bulgarian is to avoid the realization of this participant in the direct object position in contrast to English, as demonstrated by the results in the studies. This is considered as a reflection of the general asymmetry displayed between English and Bulgarian (discussed in section 1.2.2). Where English has a tendency to use intransitive verbs as transitive, thus making efficient use of the direct object position, the tendency in Bulgarian goes in the opposite direction, as transitive verbs are commonly used intransitively, thus avoiding the overt expression of a direct object.

Continuations regarded as *Absorber* (recognized by the overtly realized *Path* component) were received for 12% of the English sentences and for a considerably fewer sentences in Bulgarian (only 2,5%). These results are illustrated in the examples in (20) below.

- (20) a. Bob kicked *the can* into the river.
- b. Lili pljasna *kartite* na masata.
Lilly slapped cards-the on table-the
- b' Lilly slapped the cards on the table.

The value of *Creation* was given to an element perceived as created in the process of or as a result of the event denoted by the head verb in the sentence, either in the *Concrete Physical Impact* sense or in an *Extended Meaning* sense of the verb at hand as illustrated in the examples in (21) below.

- (21) a. Tom tapped *a tune* on the bar. (*Concrete Physical Impact*)
- b. Frank struck up *a song* in the pub. (*Extended Meaning*)
- c. Lili pljasna *šamar* na Ivan. (*Concrete Physical Impact*)
Lilly slapped slap on Ivan
- c' Lilly slapped Ivan (on the face).

Such continuations were received for only 3% of the Bulgarian sentences (only in *Concrete Physical Impact* sense) and for 2% in the English.

In addition, 10% and 11% of the stimuli (in English and Bulgarian, respectively) elicited an element specifying *Manner*, as illustrated in the examples in (22) below.

- (22) a. Tom tapped *asynchronously*.
 b. Lili pljasna *šumno* s rüce.
 Lilly slapped loudly with hands
 b' Lilly slapped loudly with her hands.

Many of the *Manner* phrases, however, were elicited in continuations provided for iterative events, which was expected since the presence of *Iteration* was sufficient for the implication of the element in *Contact* relation which might not be realized overtly (as already discussed and demonstrated with the examples in (14) above).

5.3.1.2 Results for stimuli of the type *Subject*_[Initiator] *Verb Object* ____

Compared to the sentences discussed so far, the stimuli in this group¹¹ included also an object, thus they were virtually "completed," i.e. they were grammatical and no obligatory continuation had to be provided. Yet, only one English sentence and 5 Bulgarian sentences did not receive a continuation. These were respectively 0,6% and 3% of the sentences for this group of verbs. Another ten English and four Bulgarian sentences elicited a continuation introduced with a conjunction, i.e. not directly related to the semantics of the verb at hand.

According to the lexical representation given in Chapter 4, these verbs specified for *Source Extension* and an eventual *Place of Contact* and namely this information was expected to get activated upon recognition of the verbs and subsequently to appear in a continuation provided by the subjects in the tests.

¹¹ The results for this type of stimuli are displayed in Table 2 in Appendices D and E, for English and Bulgarian, respectively.

The results demonstrated a very high preference for this type of stimuli to be continued with phrases denoting a *Source Extension*. More than half of the sentences in Bulgarian (56%) and 30% of the English sentences elicited a continuation realizing a participant with this value, as illustrated in the examples in (23) below.

- (23) a. Mark hit the door *with his fist*.
 b. Straxil promuši čoveka *s nož*.
 Strahil stabbed human-the with knife
 b' Strahil stabbed the man with a knife.

As mentioned above, this was exactly what I expected in line with the predictions of the model since the stimuli were already complete sentences.

It should also be noticed that neither the English nor the Bulgarian stimuli for *kick (ritna)* elicited a continuation with the value of *Source Extension*. This was again predicted by the model, as *kick* lexicalizes for a specific *Source Extension* and the overt syntactic realization of this participant would, in most of the cases, yield redundancy (cf. the discussion in section 3.3.4 of this work).

A continuation with the value of *Place of Contact* was given to 17% of the English sentences and 14% of the Bulgarian. However, these were concentrated among the answers for *kick (ritna)* and *stab (proboda/promuša)*, as illustrated in the examples in (24) and (25) below.

- (24) a. Harry kicked the man *in the knee*.
 b. Brian stabbed the man *in the chest*.

- (25) a. Borjana ritna muža *v kraka*.
 Borjana kicked man-the in leg-the
 a' Borjana kicked the man in the leg.
 b. Straxil promuši čoveka *v sŭrceto*.
 Strahil stabbed human-the in heart-the
 Strahil stabbed the man in the heart.

This was expected as it reflects the semantic features of the direct object, which is [+animate] for *kick* (*ritna*) and *stab* (*promuša*). *Place of Contact* of non-animate objects is rarely realized, as it does not convey significant information in comparison to *Place of Contact* of an animate participant. Additionally, this is demonstrated in the difference of the results for the Bulgarian stimuli with *stab* (*proboda/promuša*), where the stimulus containing the non-animate direct object received only one continuation (given in the example in (26) below) with the value of *Place of Contact*, which specifies multiple contact points.

- (26) Bobi probode mesoto s nož na *njakolko mesta*.
 Bobby stabbed meat-the with knife on several places
 Bobby stabbed the meat *in several places* with a knife.

A considerably large number of the continuations for the stimuli of the type *Subject*_[Initiator] *Verb Object*___ referred to *Manner* specification. These were 25% of the English sentences and 14% of the sentences in the Bulgarian test. This high rate could be explained with the presence of a motion component in the event lexicalized by the verbs at hand. Thus, many of these phrases specified this motion component, as illustrated in the examples in (27) below.

- (27) a. Jack tapped the table rapidly.
 b. Stojan udari vratata silno/leko.
 Stojan hit door-the hard/lightly
 b' Stojan hit the door hard/lightly.

A simple comparison between the quantities of *Manner* phrases elicited by the verbs discussed and a group of distracter verbs (verbs of experience and perception) reveal a considerable drop in percentage – only 6% of the sentences received a continuation with a *Manner* value.

Continuations with other values (*Time*, *Reason/Cause*, *Quantification*, and *Other*) were given to 21% of the English sentences and 14% of the Bulgarian sentences.

5.3.1.3 Results for stimuli of the type *Subject Verb* ____

As mentioned above, the last group of stimuli denoted non-canonical situations lexicalized by the verbs at hand, where the subject could not be assigned the value of *Initiator*, but carried the value of *Source*, *Source Extension*, or *Absorber*.

Since all the verbs in this group denote a situation where *Contact* is attained and only one of the participants in this situation was overtly realized in the stimuli, the primary expectations were for continuations referring to the second participant in the situation, i.e. a *Criterial* element by which the situation is identified as belonging to this type. Thus, I have expected mainly continuations with the value of *Limit*.

The results from the tests (given in Table 3 in Appendices D and E for English and Bulgarian) firmly confirmed my expectations. 31% and 45% of the sentences, respectively for English and Bulgarian, were continued with a phrase in the direct object position carrying the value *Limit*. Another 29% and 31%, respectively, received a continuation regarded as *Limit*, but realized as a prepositional phrase. Thus, 60% of the English sentences and 76% of the Bulgarian sentences elicited a continuation marked as *Limit*, as illustrated in the examples in (28) and (29) below.

- (28) a. The ball hit *the window*.
 b. The rain hammered *on the roof*.
- (29) a. Topčeto udari *zemjata*.
 ball-the hit ground-the
 a' The ball hit the ground.
- b. Džždüt čukaše *po prozoreca*.
 rain-the hammered on window-the
 b' The rain hammered on the window.

In addition, an *End of Path* component was elicited in 40% of the results for the English stimulus "*The book slapped* ____" and 60% of the results for its Bulgarian correlate "*Knigata pljasna* ____," as illustrated in the examples in (30) below.

- (30) a. The book slapped *on the table*.
 b. Knigata pljasna *na poda*.
 book-the slapped on floor
 b' The book slapped on the floor.

Besides, 33% of these results for the English stimulus "*The book slapped ____*" included a continuation marked as *Result*, as exemplified in (31).

- (31) The book slapped *shut*.

Continuations related to *Manner* were given for only 9% of the English sentences, spread evenly across all the stimuli. For Bulgarian, there were 10% of *Manner* phrases elicited mainly for the two stimuli with *pljasna (slap)*.

It is also interesting to mention the extremely high percentage of continuations specifying *Time* in the English test – 15% of all the sentences. However, most of them (11%) were given for the verb *strike*, as illustrated in the example in (32) below.

- (32) The disaster struck *at midnight*.

This find could be explained with the specificity of the stimulus (lexicalizing a metaphorical situation) and the extended meaning of the verb *strike* when combined with the lexical item at hand, as for example *disaster*. In this case, the verb can be considered an element of a set collocation as discussed in Hoey (2005), where on the basis of corpus data it is suggested that each word primes another word or words which are part of its typical collocation. This is in line with the connectionist models approach discussed also in Koenig et al. (2002).

Furthermore, in Jackendoff (2002, among others) such expressions are considered as idioms or semi-idioms. These are regarded as complex conceptual representations which are stored in long-term memory and filled in with variables when used in language production. On this account, these expressions are similar to constructions in the sense of Construction Grammar approach, discussed in section 2.2.

Only 7% of the English sentences elicited other continuations (not discussed above) and even fewer of the sentences in Bulgarian – barely 3%.

Finally, the sentences, which did not elicit a continuation or the continuation received was introduced with a conjunction, constituted slightly more than 1% of the English sentences and 4% of the sentences in Bulgarian.

5.3.2 Results for Verbs Denoting a Conditioned Event

The second group of verbs consists of verbs denoting a *Conditioned Event* (cf. the discussion in section 4.2.2 of this work.) The verbs discussed in this section are the English *throw*, *cast*, *shoot*, and *cut* and the Bulgarian *metna*, *xvŭrlja*, *streljam*, and *reža*, respectively, as well as the Bulgarian *sŭborja* (*knock down/over*) and *odraskam* (*scratch*). In addition, some verbs displaying dual lexicalization pattern, like the English *knock*, *hammer*, *scratch*, *smash*, and *dab* and the Bulgarian *blŭsna* (*knock*, *push*), (*za*)*čukam* (*hammer*) are also included in the discussion.

The expectations were mainly for continuations denoting participants present in the formal representations of the verbs at hand (cf. the discussion in section 4.2.2), i.e. lexically encoded participants. Thus, I have expected continuations referring to *Absorber* (if not included in the stimuli), *Path* components (mainly for the verbs *xvŭrlja* (*throw*), *metna* (*cast*), and *sŭborja* (*knock down/over*)), *Limit* (for verbs displaying the dual lexicalization pattern, as well as for *shoot*, which encodes both *Absorber* and *Limit*), *Place of Contact* (again for verbs displaying the dual lexicalization pattern), and *Source Extension*.

5.3.2.1 Results for stimuli of the type *Subject*_[Initiator] *Verb* ____

As expected, most of the continuations received for the stimuli of type *Subject*_[Initiator] *Verb* ____ referred to participants with the value *Absorber* realized in the direct object position, as illustrated in the examples in (33) below.

- (33) a. Margaret cut *the tomatoes*.

- b. Maria otrjaza *kabela*.
 Maria off-cut cable-the
 b' Maria cut off the cable.

These were 60% of the English sentences for the verbs *throw*, *cast*, *shoot*, and *cut* and 75% of the Bulgarian sentences for the verbs *metna*, *xvŕlja*, *streljam*, *reža*, *sŭborja*, and *odraskam*.

Besides, continuations with the value of *Absorber* were given to 45% of the English sentences with verbs of dual lexicalization pattern (*knock*, *hammer*, *scratch*, *smash*, and *dab*) and 50% of those in Bulgarian (*blŭsna*, *začukam*).

As predicted by the model, the same verbs, elicited also continuations marked as *Limit* - 43% for the five English verbs and 55% for the two Bulgarian verbs.

Illustrative examples of the various types of continuations are presented in (34) and (35) below.

- (34) a. Terry hammered *the nail*. (Absorber)
 b. Terry hammered *on the door*. (Limit, PP)
 c. Steven dabbled *his brow*. (Limit, O_d)
 d. Steven dabbled *oil on a canvas*. (Absorber + Limit)
- (35) a. Todor začuka *pirona*. (Absorber)
 Todor pref-hammered nail-the
 a' Todor hammered the nail.
- b. Todor začuka *po vratata*. (Limit, PP)
 Todor pref-hammered on door-the
 b' Todor began hammering on the door.
- c. Jordan blŭsna *Maria bez da iska*. (Limit, O_d)
 Jordan knocked Maria without to want
 c' Jordan bumped into Maria accidentally.

- d. Todor začuka *pirona v stenata*. (Absorber + Limit)
 Todor pref-hammered nail-the in wall-the
 d' Todor hammered the nail *into the wall*.

It should be noticed that both the Bulgarian and the English verb *hammer*, as well as the English *dab*, allow for simultaneous expression of participants with value of *Absorber* and *Limit* as illustrated in (34d) and (35d). However, the status of the participant with the value of *Limit* in such situations is open for discussion since in the limiting case, the participant which can be described as the last entity of the *Force* arc, i.e. the *Limit*, can alternatively, in the *Monodevelopment* dimension, be conceptualized as marking the *End* of the *Trajectory Line* of the *Mover*, that is, carrying the value of *End of Path*. Therefore, this participant can be characterized by the set of features [*Limit_i*, *End of Path_i*], where its index *i* will also appear in *Contiguous to: i*, thus making a reference to the entity where the *Conditioning* stops. However, for the time being I will continue to refer to this participant simply as marked with the value of *Limit*.

In addition, it must be mentioned here that the Bulgarian prefix *za-* can be related to different interpretations,¹² according to the situation at hand. Thus, if the verb is used to denote a situation of *Contact* only, as in the example in (36b), the single interpretation of the prefix would be to indicate of the start of the process denoted by the verb, as illustrated in the English translation in (36b'). However, if the verb is used to lexicalize a conditioned event, as in the sentence in (36a), the only possible interpretation is a perfective reading, i.e. it indicates a completed process.

The *Place of Contact* value could be assigned to 17% of the continuations given for the English verbs *knock*, *hammer*, *scratch*, *smash*, and *dab* and 2% of the Bulgarian verbs *blūsna* and *začukam*, as illustrated in the examples in (36) below.

- (36) a. John scratched *his nose*.
 b. John scratched an itch *on his arm*.
 c. Ivan odraska rŭkata si.
 Ivan scratched hand-the refl.poss.cl.

¹² Cf. Guentcheva (2002) for a discussion on the semantics and the functions of prefixes in Bulgarian.

- c' Ivan_i scratched his_i hand. (on the reading that he hurt his hand)
- d. Ivan odraska Pešo po rŭkata.
Ivan scratched Pešo on hand-the
- d' Ivan scratched Pešo on the hand.

In addition, the English verbs *cut* and *shoot* also elicited continuations regarded as *Place of Contact*, as illustrated in the examples in (37) below.

- (37) a. Margaret cut her knee.
b. Tom shot his own foot.
c. Tom shot his friend in the foot.

Much as expected, there were also continuations referring to *Path* components with the verbs *xvŭrlja* (*throw*), *metna* (*cast*), and *sŭborja* (*knock down/over*), as illustrated in the examples in (38) below.

- (38) a. The girl threw her doll *on the floor*
b. Momičeto xvŭrli moneta *vŭv fontana*.
girl-the threw coin in fountain-the
b' The girl threw a coin in the fountain.

These appeared in 19% of the English sentences and 19% of the Bulgarian sentences with verbs denoting a conditioned event only.

In addition the verb *shoot* was used to denote some metaphorical situations where a *Path* component was overtly realized, as illustrated in the examples in (39) below.

- (39) a. Tom shot *out the door*.
b. Tom shot *forward*.

These are non-canonical situations where the verb *shoot* is used as a *Manner of Motion* verb. This find, however, is not unexpected as it relates to the possibility of the participant with the set of values [*Absorber, Mover*] to be realized in the subject position in the *Concrete Physical Impact* sense of the verb. Similar situations were discussed in section 4.2.2.3 and illustrated in an example repeated in (40) below.

- (40) a. Topkata se izstrelja navisoko.
 ball-the refl.cl. shot high up
 a' The ball shot high up.

Thus, the results illustrated in the examples in (39) above were also predicted by the model as possible syntactic patterns to be displayed by the verb *shoot*, although lexicalizing non-canonical situations.

The verbs of dual lexicalization pattern also elicited continuations referring to *Path* components, which were present in 3 of the completions for the Bulgarian verb *blūsna* (*push*) and 2 for each of the verbs *knock* and *smash*, as illustrated in the examples in (41) below.

- (41) a. David knocked over a glass *on the floor*.
 b. Jordan blūsna Gošo *po stūlbite*.
 Jordan pushed Gošo on stairs
 b' Jordan pushed Gošo down/up the stairs.

Finally, continuations referring to a participant with the value of *Source Extension* were given to 8% of the English sentences and 5% of the Bulgarian sentences with all the verbs discussed in this section, as illustrated in the examples in (42) and (43) below.

- (42) a. Tom shot *the gun*.
 b. Margaret cut the long grass *with the scythe*.
 c. Steven dabbed *his paintbrush* in the paint.
 d. Steven dabbed his wounded arm *with a piece of tissue*.

- (43) a. Toni streljaše *s pistol*.
 Toni shoot with gun
 a' Toni was shooting with a gun.
- b. Jordan blūsna *s jumruk* po masata ...
 Jordan knocked with fist on table-the
 b' Jordan knocked his fist against the table.

Continuations which did not refer to lexically encoded semantic participants were given to 15% of the English sentences and 12% of the Bulgarian sentences for all the verbs discussed in this section.

5.3.2.2 Results for stimuli of the type *Subject*_[Initiator] *Verb Object*_[Absorber] _____

As mentioned above, this kind of stimuli were syntactically of the type *Subject-Verb-Object*. In addition, all the participants in the object position could be assigned the value of *Absorber*. Thus, the verbs were meant to denote situations including a *Conditioned* event (even those verbs that could lexicalize events where only *Contact* was attained).

Therefore, I expected a high percentage of continuations related to participants included in the outlined lexical representations of the verbs as lexicalizing a *Conditioned* event (cf. the discussion in section 4.2.2). That is, continuations referring mainly to participants with the values of *Source Extension* and *Path* (*length, origin, end, orientation*). Besides, a number of continuations related to *Limit* and *Place of Contact* were also expected (only if encoded in the verb at hand, e.g. for *shoot*).

In addition, since the stimuli were all grammatical sentences from the beginning, I expected relatively higher percentages of uncompleted sentences, or sentences continued with phrases referring to information which is not lexically encoded in the verbs at hand.

The overall results showed that, in fact, a significant percentage of the continuations related to participant information included in the semantic representations of the verbs discussed in section 4.2.2.

Thus, continuations referring to *Source Extension* were given to 9% of the English sentences and 10% of the Bulgarian sentences, as illustrated in the examples in (44) below.

- (44) a. Nick smashed the mirror *with his fist*.
 b. Toj otrjaza dŭrvoto *s trion*.
 he off-cut tree-the with saw
 b' He cut off the tree with a saw.

These were spread rather even across the verbs, if we take into account that the verbs *throw*, *cast*, and *hammer* were not expected to elicit almost any continuations with that value (cf. the discussion in section 3.3.4).

21% of the English sentences and 20% of the Bulgarian sentences were completed with a phrase referring to a *Path* component, as illustrated in the examples in (45) below.

- (45) a. Susan threw the ball *over the hedge*.
 b. Sonja xvŭrli topkata *ot prozoreca*.
 Sonja threw ball-the from window-the
 b' Sonja threw the ball from the window.

As expected, a *Path* component was elicited mainly for the verbs *throw* (*xvŭrlja*) and *cast* (*metna*), since the participant expressed as the direct object carries also the value of *Mover*. However, the stimuli with the verb *knock* (*blŭsna*) also elicited continuations that related to a *Path* element, as illustrated in the examples in (46) below.

- (46) a. Mary knocked the chair *off the roof*.
 b. Marija blŭsna stola *na zemjata*.
 Maria knocked chair-the on ground-the
 b' Maria knocked the chair onto the ground.

In addition, 14% of the English sentences and 22% of the Bulgarian sentences received continuations that could be assigned the value of *Limit*, as illustrated in the examples in (47) and (48) below.

- (47) a. Mary knocked the chair *into the wall*.
 b. Carry scratched her leg *on the edge of the chair*.

- (48) a. Marija blūsna stola *v stenata*.
 Maria knocked chair-the in wall-the
 a' Maria knocked the chair into the wall.
 b. Tja odraska kraka si *na ogradata*.
 she scratched leg refl.poss.cl. on fence-the
 b' She_i scratched her_i leg on the fence.

Most of the completions, however, were given for the verb *hammer* (*začuka*) as illustrated earlier in the examples in (36d) and (36d'). However, as discussed earlier in section 5.3.2.1 for the verbs *hammer* and *dab*, the status of the participants marked as *Limit* in situations including a *Conditioned* event is open for discussion.

The variation in the possible continuations for *knock* (*blūsna*), either related to a *Path* component or to a *Limit*, can be related to the ability of the verb to choose between the two lexicalization patterns (denoting a *Contact* situation or a *Conditioned* event). This could also be said for the verb *scratch* (*draskam*) which received rather similar continuations, yet with different semantic values,¹³ as illustrated in the examples in (49) below.

- (49) a. Carry scratched her leg *with a rusty nail*. (*Source Extension*)
 b. Carry scratched her leg *on the tree*. (*Limit*)

¹³ Cf. section 4.2.2.2 for analysis and discussion.

Besides, a *Place of Contact* value could be assigned to almost 10% of the English sentences and slightly more than 10% of the Bulgarian sentences, as illustrated in the examples in (50) below.

- (50) a. The man shot the thief *in the leg*.
 b. Čovekūt prostrelja kradeca *v kraka*.
 man-the shot thief-the in leg-the
 b' The man shot the thief in the leg.

As predicted by the model, these were given almost exclusively for the verb *shoot*. Thus, only two other stimuli elicited participants with the value *Place of Contact* (only 3,7% of the sentences) and these were with the English verbs *dab* and *scratch*, when conceptualized as denoting a *Contact* situation.

With respect to this, it is interesting to mention that although the stimuli contained a direct object, meant with the value of *Absorber*, some sentences with verbs of dual lexicalization pattern still elicited continuations related to the use of verb as denoting an event where only *Contact* was implicated, as illustrated in the example in (51) below.

- (51) Carry scratched her leg *where the mosquito had bitten her*.

Thus, the continuation in (51) above (which could only be assigned the value of *Place of Contact*) forced the reading of the sentence into denoting solely a *Contact* situation.

Altogether 22% of the English sentences¹⁴ and 25% of the Bulgarian sentences received continuations specifying *Manner*, *Time*, *Reason/Cause*, *Quantification*, and *Other*. This was even less than originally expected, thus providing more positive evidence that the information which was thought to be present in the lexical

¹⁴ The percentage for English does not include the numbers under the column *Other*, as these are predominantly satellite particles which are believed to reinforce and further specify the meaning of the verb at hand (cf. the discussion in the introductory chapter, section 1.2.1).

representation of the verbs at hand is indeed accessed upon recognition of the verbs during the online continuation studies.

Finally, only 8% of the continuations for the English sentences and 14% of the continuations for the Bulgarian sentences fall into the columns *Conjunction* and *No Continuation*.

5.3.2.3 Results for stimuli of the type *Subject Verb* ____

The last group of results to be discussed were elicited for stimuli of the type *Subject Verb* ____ denoting non-canonical situations lexicalized by the verbs at hand.

The expectations varied according to the information carried by the individual verbs and the range of semantic interpretations of the stimuli. To what extent the expected continuations coincided with the actual results received is discussed together with the results for each particular type of continuation.

34% of the English sentences and 17% of the Bulgarian sentences received a continuation referring to a participant with the value *Absorber*, as illustrated in the examples in (52) below.

- (52) a. The knife cut *the carrot*.
 b. The machine cast out *oil*.
 c. The gun shot *a bullet*.
- d. Nožūt režeše *tvŕrdia xľjab*.
 knife-the cut hard-the bread
 d' The knife was cutting the hard bread.
- e. Mašinata mjataše *testo vŭv furnata*.
 machine-the cast dough in oven-the
 e' The machine was casting dough into the oven.

- f. Puškata streljaše s xalosni patroni.
rifle-the shot with blank cartridges
- f' The rifle was shooting blanks.

This type of continuation was expected with the verbs *cut* (*reža*) and *cast* (*metna*), as well as with *shoot*¹⁵ (*streljam*). However, the English stimuli for *knock* and *scratch* also elicited a number of continuations (50% of the sentences with *knock* and 17% of the sentences with *scratch*) referring to a participant with the value *Absorber*, as illustrated in the examples in (53) below.

- (53) a. The chair knocked over the little girl.
b. The chair knocked the vase off when it fell.
c. This surface scratches my skin.
d. This surface scratches softer surfaces.

The results demonstrate that subjects in these stimuli were not perceived exclusively as participants performing a *Monodevelopment* (as intended initially and which was the only reading of the Bulgarian example) but they could also be regarded as a *Source* (which was impossible in Bulgarian due to the anaphoric clitic *se*). This is in line with the predictions of the model, as both verbs are believed to display the dual lexicalization pattern.

Also in accordance with the expectations, the value of *Limit* was assigned to 18% of the English sentences and 14% of the Bulgarian sentences, as illustrated in the examples in (54) below.

- (54) a. The gun shot *the man*.
b. The glass smashed *against the wall*.
c. Stolūt se blūsna v *stenata*.
chair-the refl.cl. knocked in wall-the
c' The chair knocked into the wall.

¹⁵ In English, the stimulus with *Concrete Physical Impact* sense of the verb.

As already mentioned, some of the stimuli in Bulgarian contained the anaphoric clitic *se*, thus leaving only one option for the syntactic realization of the participant marked as *Limit*. That is, it could be overtly realized only as a prepositional phrase (cf. also the discussion in the introductory chapter, section 1.2.3).

Continuations related to a *Path* component were expected mainly for the English verbs *cast* and *shoot* (in its extensive reading) and for the respective Bulgarian verbs *metna* and (*se*) *izstreljam*, as well as for the Bulgarian *sǔborja* (*knock down/over*).

Altogether, 27% of the English sentences and 20% of the Bulgarian sentences received a continuation related to different *Path* components, as illustrated in the examples in (55) and (56) below.

- (55) a. The machine cast its motor *into the sea*. (*Concrete Physical Impact*)
 b. The machine cast a large shadow *on the factory floor*. (*Extended Meaning*)
 c. The car shot *out of the drive*. (*Path Origin*)
 d. The car shot straight *into the grand oak tree*. (*Path End*)
 e. The car shot down the street. (*Path Orientation*)
 f. The car shot *through the traffic lights*. (*Path Length*)

- (56) a. Mašinata mjataše testo *vǔv furnata*. (*Concrete Physical Impact*)
 machine-the cast dough in oven-the
 a' The machine was casting dough into the oven.
 b. Mašinata mjataše kupišta informacia *kǔm printera*. (*Ext. Meaning*)
 machine-the cast piles information towards printer-the
 b' The machine was sending a great deal of information towards the printer.
 c. Kolata se izstrelja *ot tunela*. (*Path Origin*)
 car-the refl.cl. shot from tunnel-the
 c' The car shot from the tunnel.

- d. Kolata se izstrelja *v propastta*. (Path End)
 car-the refl.cl. shot in precipice-the
 d' The car shot into the precipice.
- e. Kolata se izstrelja *kŭm moreto*. (Path Orientation)
 car-the refl.cl. shot towards sea-the
 e' The car shot towards the seaside.
- f. Kolata se izstrelja *po pistata*. (Path Length)
 car-the refl.cl. shot along track-the
 f' The car shot along the racing track.

As illustrated in the examples above, a participant specifying *Path* could appear both in situations denoted by the *Concrete Physical Impact* sense of the verb at hand and in sentences where the verb is used with *Extended Meaning*.

Outside the widespread expectations, completions which could be regarded as specifying *Path* were elicited also by the English verb *cut*, as illustrated in the examples in (57) below.

- (57) a. The knife cut *through the bread*.
 b. The knife cut *into her flesh*.

In line with the proposed model, these continuations could be explained with the semantic values of the participant realized as the subject of the sentence. Thus, *the knife* is assigned on the different dimensions the set of values [*Source Extension, Mover*]. Being a *Mover*, this participant performs a *Monodevelopment* with *Medium: Location* and therefore its *Trajectory* line could be specified as well. Thus, these continuations could also be predicted by the lexical representation of the verb *cut* as outlined and discussed in section 4.2.2.2.

In Bulgarian, this could be achieved only with the means of prefixes which modify and further specify the meaning of the verb, as illustrated in the example in (58) which is a semantic equivalent of the English example in (57b) above.

- (58) a. Nožūt se vřjaza v plūtta i.
 knife-the refl.cl. pref-cut in flesh her
 a' The knife cut into her flesh.

However, in the example in (58), the verb *se vřjaza* (*cut into*) is used as a *Manner of Motion* verb like other transitive verbs used to lexicalize intransitive situations discussed already in section 4.2.3.

Finally, this group of stimuli received in general a very high percentage of continuations specifying *Manner* – 19% of the sentences in English and 44% of the sentences in Bulgarian. However, these were elicited mainly for two of the English stimuli and three of the Bulgarian, as illustrated in the examples in (59) and (60) below.

- (59) a. This surface scratches *easily*.
 b. The knife cut *easily* into the meat.

- (60) a. Tazi povŭrxnost se draska *lesno*.
 this surface refl. cl. scratch easily
 a' This surface scratches easily.

- b. Nožūt režeše *dobre*.
 knife-the cut-past.progr. well
 b' The knife was cutting well.

- c. Puškata streljaše *točno*.
 rifle-the shoot-past.progr. accurately
 c' The rifle was shooting accurately.

Also known as the *Middle Alternation* (Levin 1993), the sentence in (59a) and the corresponding Bulgarian one in (60a) was expected to receive many completions referring to *Manner*, which is a characteristic feature of this alternation together with *The Simple Present Tense* form of the verb.

However, I wanted to check whether a change in the aspectuality of the verb would skew the results in some direction. Thus, I expected that a change from habitual reading of the verb into a perfective reading would elicit more continuations related to lexically encoded participant information. This could not be done straightforwardly in Bulgarian, as a change in the aspectuality of the verb would also trigger a change in the syntactic pattern – a perfective prefix would elicit an obligatory direct object, as illustrated in the example in (61) below.

- (61) a. Nožūt otrjza *(hljaba).
 knife-the pref-cut bread-the
 a' The knife cut off the bread.

Therefore, I constructed the English stimulus for *cut* in the *Past Simple Tense*. A comparison of the results for the two stimuli presented in (59) above showed on the one hand a drastic drop of the completions referring to *Manner* specification – from 57% for the stimulus with *scratch* to 23% for the stimulus with *cut*. On the other hand, an increase of the continuations related to a participant with the value Absorber was demonstrated – from 17% for the stimulus with *scratch* to 47% for the stimulus with *cut*. Besides, most of the continuations referring to the value of *Manner* with the verb *cut* appeared together with another phrase, as exemplified in (59b) above.

5.3 Conclusions

This chapter presented the results of the online sentence continuation studies conducted as part of my research project. The studies aimed at providing experimental evidence for or against the adopted model of lexical representation of verbs across languages. Together with a short description of the design of the tests and the methodology used in the studies, I presented the results organized into groups according to the types of verbs analyzed and the variety of stimuli used in the tests. Special attention was paid to continuations related to lexically encoded participants as included in the semantic representations suggested for the verbs in the research.

The overall finds revealed positive evidence for the activation and the use of lexically encoded information in the online sentence processing and language production. Thus, the results confirmed the predictions of the model and the analyses of the verbs discussed in Chapter 4.

The continuations used by native speakers substantiated the grouping of the verbs according to the two types of conceptual representations suggested in sections 4.2.1 and 4.2.2 respectively. The *Limit–Absorber(Monodeveloper)* dichotomy outlined in the lexical representations of the examined verbs was reflected in the overt realizations of the respective participants as expected and predicted by the model.

Besides, the ambiguity in the perception of some sentences (with *knock* and *scratch*, for example) confirmed the hypothesis that some verbs may choose between the two representational formats. Thus, the net-like grouping of verbs suggested in Chapter 4 corresponds to the evidence provided by native speakers in the online studies.

In addition, the high percentages of continuations related to participants with the values of *Source Extension* and *Path* also supported the expectations of the model that participants with these values were part of the lexical representations of the verbs as outlined in the previous chapter.

Furthermore, the empirical evidence in the continuations provided by native speakers in the online studies supported also the multi-dimensional model of lexical representation of verbs discussed in section 3.3 and outlined in detail for the examined verbs in Chapter 4. Thus, each participant was not labelled with a single semantic role but specified by a set of co-indexed values on the different dimensions which reflected its involvement in the various aspects of the situation lexicalized by the verb at hand.

Finally, an emphasis must be set on the necessity of follow-up empirical tests assessing the details in the representations of correlate verbs to account for the subtle differences in event processing and mapping of conceptual participants into lexical items and patterns of grammaticalization across languages.

6. Concluding Remarks

In the course of this project, I have investigated the information that could be encoded in the lexical representation of verbs and the mapping of semantic participants onto lexical items and various syntactic patterns across languages.

Much in the tradition of the current linguistic approaches discussed in Chapter 2 of this work, I have explored various linguistic phenomena rising on the interface of conceptual structure with syntax. However, my work concentrated on the evidence provided by empirical data and native speakers' intuition, expressed overtly through tasks involving sentence processing and language production, used in assessing the information believed to be lexically encoded in verbs. This approach was in line with recent studies conducted independently by two research teams discussed in Chapter 3.

For the purposes of the project, empirical data from two Indo-European languages, English and Bulgarian, were analysed for the type of semantic participants involved in the situations lexicalized by a set of verbs and the possible syntactic realizations of these participants. Each of the participants in the situations denoted by the verbs was ascribed a bundle of semantic features characterizing it on different dimensions which reflected the various aspects of involvement of the participant in the situation at hand. This representational format followed a framework called *The Sign Model* (in Dimitrova-Vulchanova 1996/99) presented in section 3.3 of this work.

The corpus data analyses discussed in Chapter 4 provided evidence for a distinction between two basic types of situations lexicalized by the verbs at hand and dubbed *Contact* (situation) and (a situation of) *Conditioning* or *Conditioned* event.

The examined verbs displayed a strong tendency to group according to the type of situation they can lexicalize. Thus, verbs denoting *Contact* situations shared patterns of alternation as opposed to verbs denoting a *Conditioned* event. Hence the verbs were grouped according to the demonstrated similarities in their syntactic behaviour which was directly related to the semantic features ascribed to the participants involved in the situations lexicalized by the verbs at hand.

In addition, both languages attested the employment of the *Dual Lexicalization Pattern*, i.e. some verbs were used to lexicalize events of both types. Thus, these verbs can choose a frame (representational format) according to the type of situation they lexicalize. Therefore, the verbs could not be merely enumerated in various lists. Instead, a more net-like pattern of distribution was employed, accounting for the possibility of one verb to lexicalize situations of different types. Thus, the verbs were linked to each other and grouped in accordance with the types of situations they can lexicalize and the set of values of their participants.

The information extracted from the corpus data analyses and the respective lexical representations outlined for the verbs at hand were then checked against the native speakers' intuition, as discussed in Chapter 5. Following the adopted model of lexical encoding of verbs I have expected higher percentages of continuation related to participants included in the verbs' suggested representations. These expectations were based on recent research conducted by Koenig et al. (2002, 2003) showing that lexically encoded participant information is activated upon recognition of the word and is more likely to appear in a following language production task.

The results received in my online sentence continuation studies unambiguously confirmed the predictions I made in advance and thus substantiated the underlying truthfulness of the outlined representational format. Thus, I received higher percentages of continuation related to participants believed to be lexically encoded in the verbs at hand. In addition, there were relatively many continuations substantiating the grouping of the verbs outlined in Chapter 4 and the multi-dimensional model of representation.

However, further empirical evidence is necessary to elaborate on the representational model adopted including a larger variety of verbs, which could be grouped according to that model and distributed in a VerbNet – a network of interconnected verbs within a single language, as well as across languages.

Finally, the two types of situations distinguished in the empirical data are remarkably parallel (and may be implicitly related) to the conceptual analysis of time discussed in section 2.4.3.3 (cf. Jackendoff (1987); Nikanne (1995)). Thus, I would compare the pure *Contact* situation to a point (in time), while the *Conditioning* event can be represented as a region (a monodevelopment line). There are verbs that can lexicalize only one of these types (for example *touch* vs. *cut*). However, many verbs may lexicalize each of the two. These are cases when a point borders a region, i.e. a *Contact* situation may be conceptualized as preceded or followed by a *Conditioned* event. Evidence for such cases was encountered the corpus data where verbs that usually denote a situation of contact may be further employed to lexicalize a conditioned event or the other way round. Thus, verbs like *scratch* may be presented as point-region (*Contact* followed by *Conditioning*), while *kick* as region-point (*Conditioning* followed by *Contact*). These finds are based on the possible syntactic patterns displayed by the verbs at hand and discussed in Chapters 4 and 5.

Although only intuitively outlined, this parallel may have its origin in the way people conceptualize the situations by chunking them into sub-events. And the two main components distinguished here are *Contact* and *Conditioning* (*Monodevelopment*), i.e. point and region. Consequently, languages differ in the combination and mapping of these components into lexical items and grammaticalization patterns which is an area open for further research and discussion.

Bibliography

- Assenova, Petja 1989/2002: *Balkansko ezikožnanie. Osnovni problemi na Balkanskija ezikov sŭjuz*. [Basic Problems of the Balkan Sprachbund]. Sofia: Nauka i izkustvo.
- Avgustinova, Tania 1997: *Word order and clitics in Bulgarian*. Saarbrücken, Universität des Saarlandes / DFKI.
- Bierwisch, Manfred 1996: How much space gets into language? In Paul Bloom, Mary A. Peterson, Lynn Nadel, and Merrill F. Garrett, editors, *Language and Space*, pages 31 – 76, Cambridge, Mass: The MIT Press.
- Borer, Hagit 2004: The grammar machine. In Alexiadou, Anagnostopoulou, & Everaert (eds) *The Unaccusativity Puzzle: Explorations of the Syntax-Lexicon Interface*. Oxford: Oxford University Press.
- Carpenter, Bob 1992: The logic of typed feature structures: with applications to unification grammars, logic programs, and constraint resolution, *Cambridge tracts in theoretical computer science* 32, Cambridge University Press.
- Carpenter, Bob 1997: *Type-logical semantics*, Cambridge, MA: The MIT Press.
- Carrier, Jill, & J. H. Randall 1992: The Argument Structure and Syntactic Structure of Resultatives. *Linguistic Inquiry* 23, pp. 173 - 234.
- Coltheart, Max 1981: The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology*, 33(A), pp. 497-505.
- Copestake, Ann 2002: *Implementing typed feature structure grammars*. CSLI lecture notes 110, Stanford: CSLI Publications.
- Copestake, Ann, A. Sanfilippo, T. Briscoe, & V. de Pavia 1993: The AQUILEX LKB: An Introduction, in T. Briscoe, V. de Pavia, & A. Copestake (eds), *Inheritance, Defaults, and the Lexicon*, Cambridge University Press, Cambridge.
- Chomsky, Noam 1957: *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam 1965: *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo 1990: *Types of A' dependencies*. Cambridge, Mass.: MIT Press
- Dabrowska, Ewa 2006: Words as Constructions, paper presented at the research seminar at the Department of Modern Languages, NTNU, Trondheim.

- Dimitrova-Vulchanova, Mila 1995: From Passive to Perfect – the Teleology of Grammar. In Moen, Inger, Hanne Gram Simonsen & Helge Lødrup (eds) *Papers from The XVth Scandinavian Conference of Linguistics*, Oslo, January 13-15.
- Dimitrova-Vulchanova, Mila 1996/99: *Verb Semantics, Diathesis and Aspect*. Doctoral dissertation, NTNU (University of Trondheim)/LINCOM, Newcastle/München.
- Dimitrova-Vulchanova, Mila 1998: A Typology of Measures, In Timo Haukioja (ed) *Papers from the 16th Scandinavian Conference of Linguistics*, Turku/Åbo November 14-16, 1996.
- Dimitrova-Vulchanova, Mila 1998a: Are Bulgarian Pronominal Clitics in the Wackernagel Position? In *Norwegian Contributions to the Twelfth International Congress of Slavists*. Cracow 1998, *Meddelelser 80*, 7-32. Universitetet i Oslo, Slavisk-Baltisk Avdeling.
- Dimitrova-Vulchanova, Mila 2000: Possessive Constructions and Possessive Clitics in English and Bulgarian DP. In Beukema, Frits & Marcel van Dikken (eds.). *Clitic Phenomena in European Languages, Linguistik Aktuell*, Amsterdam: John Benjamins, pp. 121-146.
- Dimitrova-Vulchanova, Mila 2003: On two types of results: resultatives revisited. In Beermann, Dorothee & Lars Hellan (eds) *The Proceedings of TROSS*, NTNU
- Dimitrova-Vulchanova, Mila 2004: Paths in Verbs of Motion, invited talk at *Argument Structure CASTLE Conference*, November 4-6, 2004, Tromsø University.
- Dimitrova-Vulchanova, Mila, and Lars Hellan. 1999. Clitics and Bulgarian Clause Structure. In Henk van Riemsdijk (ed.) *Clitics in the Languages of Europe*, pp. 469-514. Berlin: Mouton de Gruyter.
- Dimitrova-Vulchanova, Mila & Guliana Giusti 1999: Possessors in the Bulgarian DP. In Dimitrova-Vulchanova, M. and L. Hellan (eds.) *Topics in South Slavic Syntax*. Amsterdam: John Benjamins, pp. 163-192.
- Dimitrova-Vulchanova, Mila & Mattias Weisgerber (forthcoming): 'Concept', 'Context' and beyond: Manner of motion verbs encoding and relevant bits of information. In: Dimitrova-Vulchanova, Mila & Svetla Koeva (eds.) *Formal formats for Balkan languages*, Cambridge: Cambridge Scholars Press.

- Dimitrova-Vulchanova, Mila & Valentin Vulchanov (forthcoming) Clitic Doubling and Old Bulgarian. In: Tasmowski, L. & D. Kallulli (eds.) *Clitic Doubling in the Balkan Languages*, Amsterdam/ Philadelphia: John Benjamins.
- Dimitrova-Vulchanova, Mila & Valentin Vulchanov (in press) Old Bulgarian Syntax: the basics. In: Nikolova, S. (ed.) *Proceedings of the 25th Cyrillo-Methodian Centre Anniversary Conference*, Sofia: Bulgarian Academy of Sciences Publishing.
- Dimitrova-Vulchanova, Mila et al. (in press): Motion naming in three satellite-framed languages: a pilot study. In: Dimitrova-Vulchanova, M. & E. van der Zee (eds.) *Motion Encoding in Language*. Oxford: Oxford University Press.
- Donohue Cathryn & Mark Donohue 2004: On the Special Status of Instrumentals. In Miriam Butt & Tracey Holloway King (eds), *Proceedings of the LFG04 Conference*, CSLI Publications.
- Dowty, David R. 1979: *Word meaning and Montague grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. Dordrecht: Reidel.
- Dowty, David R. 1985: On recent analyses of the semantics of control. *Linguistics and Philosophy* 8(3), pp. 291- 331.
- Dowty, David R. 1982: Grammatical relations and Montague grammar. In Pauline Jacobson & Geoffrey K. Pullum (eds.), *The nature of syntactic representation*, Dordrecht: Reidel, pp. 79–130.
- Dowty, David R. 1986: Thematic Roles and Semantics. *Berkeley Linguistic Society* 12(3), pp. 340-54.
- Dowty, David R. 1989: On the semantic content of the notion 'thematic role'. In Barbara Partee, G. Chierchia & R. Turner (eds) *Properties, Types and Meanings*, vol. 2, 69-130. Dordrecht: Kluwer.
- Dowty, David R. 1991: Thematic proto-roles and argument selection. *Language* 67: pp. 574-619
- Fillmore, Charles J. 1968: The Case for Case. In E. Bach & R. Harms, (eds), *Universals in Linguistic Theory*, pp. 1-88. New York: Holt, Reinhart & Winston.
- Fillmore, Charles J. 1975: An Alternative to Checklist Theories of Meaning. *Berkeley Linguistic Society* 1, pp. 123-131.

- Fillmore, Charles J. 1977: The Case for Case Reopened. In P. Cole, ed., *Syntax and Semantics 8: Grammatical Relations*, pp. 59-81. New York: Academic Press.
- Fillmore, Charles J. 1985: Frames and the Semantics of Understanding. *Quaderni di Semantica* 6(2): pp. 222-53.
- Fillmore, Charles J. 1988: The Mechanisms of "Construction Grammar." *Berkeley Linguistic Society* 14, pp. 35-55.
- Fried, Miriam & Jan-Ola Östman (eds) 2004: *Construction grammar in a cross-language perspective*. In series: Constructional approaches to language, vol. 2. Amsterdam: John Benjamins.
- Goldberg, Adele 1995: *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.
- Goldsmith, J. 1976. *Autosegmental phonology*. Doctoral dissertation, MIT
- Grimshaw, Jane 1990: *Argument Structure*. Cambridge, MA: The MIT Press.
- Guentchéva, Zlatka 2002: The semantics and functions of prefixes. In: Dimitrova-Vulchanova, M., I. Krapova, D. Dyer & C. Rudin (eds.) *Balkanistica* (15) 2002, special issue on South Slavic and Balkan languages.
- Hare, Mary, Ken McRae & Jeffrey L. Elman 2003: Sense and structure: Meaning as determinant of verb subcategorization preferences. *Journal of Memory and Language* 48, pp. 281-303.
- Hellan, Lars & M. Dimitrova-Vulchanova 2000: Criteriality and Grammatical Realization. In Coopmans, Everaert and Grimshaw (eds) *Lexical Specification and insertion, CILT series 197*, pp. 165-194, John Benjamins.
- Hoey, Michael 2005: *Lexical Priming: A new theory of words and language*, London & New York: Routledge.
- Jackendoff, Ray S. 1972: *Semantic interpretation in generative grammar*. Cambridge, MA: MIT Press.
- Jackendoff, R. 1983. *Semantics and Cognition*. Cambridge, MA: MIT Press.
- Jackendoff, Ray 1987: The Status of Thematic Relations in Linguistic Theory. *Linguistic Inquiry* 18, pp. 369-412.
- Jackendoff, Ray 1990: *Semantic Structures*. Cambridge, MA: The MIT Press.
- Jackendoff, R. 1992. What is a Concept? In: *Frames, Fields and Contrast*. Lawrence Erlbaum Associates, Publishers.

- Jackendoff, Ray 1997: *The architecture of the Language Faculty*. Cambridge, MA: The Press.
- Jackendoff, Ray 2002: What's in the Lexicon? In S. Nootboom, F. Weerman, and F. Wijnen, eds., *Storage and Computation in the Language Faculty*, pp. 23-58. Dordrecht: Kluwer.
- Jespersen, Otto 1949: *A Modern English grammar: on Historical Principles*. Part 6, Syntax, Copenhagen: Munksgaard.
- Kasabov, Ivan 1990: *Semantichen rechnik – minimum*. Universitetsko izdatelstvo 'Kliment Ohridski', Sofia.
- Kay, Paul 1990: Even. *Linguistics and Philosophy* 13(1), pp. 59-112.
- Keenan, Edward 1985: Passive in the world's languages. In Shopen, Timothy (ed), *Language typology and syntactic description. Vol. 1*, pp. 243-281, Cambridge: Cambridge University Press.
- Koeva, Svetla 1998: Аргументна структура, тематични отношения и синтактична реализация на аргументите - сб. *Езиково съзнание*, София, Издателство "Наука и изкуство", pp. 206-230.
- Koeva, Svetla, Emil Doychev and Georgi Cholakov 2003: SYNTEXT - a Web-based System Designed for Frame Lexicons. In: *Proceedings from the International Workshop Balkan Language Resources and Tools*, Thessaloniki, pp. 41-48.
- Koeva, Svetla 2004: Семантично и синтактично описание на българските диалекти. In: *Българско езикознание*, том 4: *Когнитивна граматика на българския и френския език - описание и формализация*, Издателство на БАН, София, pp. 82-231.
- Koeva, Svetla 2005: Аргументи – семантични отношения и синтактична реализация. In *Аргументна структура- проблеми на простото и сложното изречение*. София, ИК "СемаРШ", pp. 25-42.
- Koeva, Svetla 2005a: Синтактични трансформации. In *Аргументна структура- проблеми на простото и сложното изречение*. София, ИК "СемаРШ", pp. 106-138.
- Koenig, Jean-Pierre 1993: Linking Constructions vs. Linking Rules: Evidence from French. *Berkeley Linguistic Society* 19, pp. 217-231.

- Koenig, Jean-Pierre Mauner, Gail Mauner, & Breton Bienvenue 2001: *Arguments for adjuncts*. Manuscript submitted for publication.
- Koenig, Jean-Pierre, Gail Mauner, & Breton Bienvenue 2002: Class Specificity and the Lexical Encoding of Participant Information. *Brain and Language* 81, pp. 224-235.
- Koenig, Jean-Pierre, Gail Mauner, & Breton Bienvenue 2003: Arguments for Adjuncts. *Cognition* 89, pp. 67-103.
- Krapova, Iliyana & Guglielmo Cinque (forthcoming). Clitic Reduplication Constructions in Bulgarian. . In: Tasmowski, L. & D. Kallulli (eds.) *Clitic Doubling in the Balkan Languages*, Amsterdam/ Philadelphia: John Benjamins.
- Kropp Dakubu, Mary Esther 2003: *Introduction to Multi-Verb Constructions in the languages of the Volta Basin*. Lecture series given at TROSS, June 23-27, 2003.
- Langacker, Ronald W. 1987: *Foundations of Cognitive Grammar* Vol. 1: Theoretical Prerequisites. Stanford: Stanford University Press.
- Langacker, Ronald W. 1990: *Concept, Image, and Symbol. The Cognitive Basis of Grammar*. Berlin: Mouton de Gruyter.
- Langacker, Ronald W. 1991: *Foundations of Cognitive Grammar* Vol. 2: *Descriptive Application*. Stanford: Stanford University Press.
- Lakoff, George 1987. *Women, Fire and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
- Levin, Beth 1993. *English Verb Classes and Alternations*. Chicago and London: University of Chicago Press.
- Levin, Beth & T. R. Rapoport 1988: Lexical Subordination, *Proceedings of CLS* 24, pp. 275-289.
- Levin, Beth & Malka Rappaport Hovav 1995: *Unaccusativity: at the syntax-lexical semantics interface*. Cambridge, MA: The MIT Press.
- Majid, Asifa et al. (forthcoming) The Semantics Categories of “Cutting and Breaking” Events across Languages. Special issue of *Cognitive Linguistics*.
- MacDonald, Maryellen C. 1993: The Interaction of Lexical and Syntactic Ambiguity, *Journal of Memory and Language*, Volume 32 (5), pp. 692-715.

- McRae K, Spivey-Knowlton, M. J. & Tanenhaus, M. K. 1998: Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language* 38, pp. 283-312.
- Meyers, Adam, Catherine Macleod, & Ralph Grishman 1996: Standardization of the Complement/Adjunct Distinction. In *Proceedings of the 7th Euralex International Congress*, Goteborg, Sweden, 1996.
- Miller, George A., Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine Miller 1990: Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography* 3(4), pp. 235-312.
- Mincheva, Angelina 1964: *Развой на дателния притежателен падеж в българския език*. София.
- Nikanne, Urpo 1990: *Zones and tiers: a study of thematic structure*. *Studia fennica linguistica* 35, Helsinki: Suomalaisen Kirjallisuuden Seura.
- Nikanne, Urpo 1995: Action tier formation and argument linking. *Studia Linguistica* 49 (1): pp. 1–32, Oxford: Blackwell.
- Nikanne, Urpo 1998: The Lexicon and the Conceptual Structure. In Timo Haukioja (ed) *Papers from the 16th Scandinavian Conference of Linguistics*, Turku/Åbo November 14-16, 1996.
- Nikanne, Urpo & Jan-Ola Östman 2006: Finland-Swedish Directionality in Conceptual Semantics and in Construction Grammar: A Methodological Dialogue. In *A Man of Measure: Festschrift in Honour of Fred Karlsson on his 60th Birthday*, a special supplement to *SKY Journal of Linguistics* 19, pp. 66-86.
- Osam, E. Kweku 2003: Introduction to the structure of Akan: its verbal and multi-verbal system, Lecture series given at TROSS, June 23-27, 2003.
- Partee, Barbara H. & Mats Rooth 1983: Generalized conjunction and type ambiguity, in R. Bäuerle, C. Schwarze, and A. von Stechow (eds.) *Meaning, Use and Interpretation of Language*, Walter de Gruyter, Berlin, pp. 361-383.
- Penchev, Jordan 2005: Словоред и интонация. In *Аргументна структура- проблеми на простото и сложното изречение*. София, ИК "СемаПШ"
- Perlmutter, David. 1978: Impersonal passives and the unaccusative hypothesis. In *BLS 4: Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society*, pp. 157-189. Berkeley, CA: Berkeley Linguistics Society.

- Petrova, Galina 2006: *Семантични роли на кратките дателни местоимения*. Изд. ДИМАНТ.
- Pollard, Carl & Ivan A. Sag 1987: *An information-based syntax and semantics. Volume 1: Fundamentals. CSLI Lecture Notes Number 13*. Leland Stanford Junior University: Center for the Study of Language and Information.
- Pustejovsky, James 1995: *The Generative Lexicon*. Cambridge, MA: The MIT Press.
- Rooth, Mats & Barbara H. Partee 1982: Conjunction, type ambiguity, and wide scope *or*, in D. Flickinger, M. Macken, & N. Wiegand (eds.) *Proceedings of the First West Coast Conference on Formal Linguistics*, Stanford Linguistics Assn., Stanford, pp. 353-362.
- Rumelhart, David E., James L. McClelland, & the PDP Research Group 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume 1: Foundations*. Cambridge, MA: The MIT Press.
- Saeed, John I. 2003: *Semantics*. 2nd Edition, In series: Introducing Linguistics 2, Malden, MA: Blackwell.
- Schütze, Carson T. 1995: PP Attachment and Argumenthood. In *Papers on Language Processing and Acquisition, MIT Working papers in Linguistics 26*, pp. 95-151.
- Schütze, Carson T. & Edward Gibson 1999: Argumenthood and English Prepositional Phrase Attachment. *Journal of Memory and Language 40*, pp. 409-431.
- Simpson, J. 1983: Resultatives. In Lori Levin, Malka Rappaport Hovav, & Annie Zaenen (eds.), *Papers in Lexical-Functional Grammar*, pp. 143-157. Bloomington: Indiana University Linguistic Club
- Spivey, M. J. & Tanenhaus, M. K. 1998: Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, pp.1521-1543.
- Talmy, Leonard 1985: Lexicalization patterns: semantic structure and the lexical forms. In Shopen, Timothy (ed.), *Language typology and syntactic description. Vol. 3 Grammatical categories and the lexicon*, pp. 57-149, Cambridge: Cambridge University Press
- Talmy, Leonard 1991: Path to realization: A typology of event conflation. *Proceedings of the Berkeley Linguistic Society 17*, pp. 480-519.

- Tenny, Carol & James Pustejovsky (eds) 2000: *Events as grammatical objects: the converging perspectives of lexical semantics and syntax*, Stanford: CSLI Publications.
- Tomić, Olga Mišeska 2004: The Balkan Sprachbund morpho-syntactic properties: An introduction. In *Balkan Syntax and Semantics*, Tomić, Olga M. (ed.), Amsterdam: John Benjamins, pp. 1–57.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, & Jan Svartvik 1985: *A Comprehensive Grammar of English Language*. London: Longman
- van der Zee, Emile & Urpo Nikanne 2000: *Cognitive interfaces: constraints on linking cognitive information*. Oxford: Oxford University Press
- von Heusinger, Klaus 2002: The Interface of Lexical Semantics and Conceptual Structure: Deverbal and Denominal Nominalizations. In: E. Lang & I. Zimmermann (eds.). *Nominalisierung*. ZAS Papers in Linguistics (ZASPiL) 27. Zentrum für Allgemeine Sprachwissenschaft (ZAS), Berlin, pp. 109-124.
- Williams, Edwin 1994: *Thematic Structure in Syntax*. Cambridge, MA: The MIT Press.
- Williams, Edwin 1987: Implicit Arguments, the Binding Theory and Control. *Natural Language and Linguistic Theory* 5, pp. 151–180.
- Wilder, Chris & Damir Čavar 1994: Word Order Variation, Verb Movement and Economy Principles, *Studia Linguistica* 48(1), pp. 46-86.

Table of Transliterations

Cyrillic (Bulgarian)		Transliteration		Sound value
Capital letters	Small letters	Capital letters	Small letters	
А	а	A	a	[a]
Б	б	B	b	[b]
В	в	V	v	[v]
Г	г	G	g	[g]
Д	д	D	d	[d]
Е	е	E	e	[e]
Ж	ж	Ž	ž	[ʒ]
З	з	Z	z	[z]
И	и	I	i	[i]
Й	й	J	j	[j]
К	к	K	k	[k]
Л	л	L	l	[l]
М	м	M	m	[m]
Н	н	N	n	[n]
О	о	O	o	[o]
П	п	P	p	[p]
Р	р	R	r	[r]
С	с	S	s	[s]
Т	т	T	t	[t]
У	у	U	u	[u]
Ф	ф	F	f	[f]
Х	х	X	x	[x]
Ц	ц	C	c	[ts]
Ч	ч	Č	č	[tʃ]
Ш	ш	Š	š	[ʃ]
Щ	щ	T	št	[ʃt]
Ъ	ъ	Ū	ŭ	[ə]
-	ь	-	j	[j]
Ю	ю	Ju	ju	[ju]
Я	я	Ja	ja	[ja]

Appendix A

An overview of the English verbs discussed (in alphabetical order) and their correlates in Bulgarian.

English	Bulgarian
cast	metna (метна)
cut	reža (режа)
drag	teglja (тегля) vlača (влача)
haul	vlača (влача)
hit	udarja (ударя)
kick	ritna (ритна)
knock (down, over)	sūborja (съборя)
knock	čukam (чукам) tropam (тропам)
pull	dūrnam (дърпам)
push	butna (бутна) blūsna (блъсна)
scratch	draskam (драскам)
shoot	streljam (стрелям)
slap	pljasna (плясна)
smash	capna (цапна)
smash	smaža (смажа)
stab	proboda (пробода), rŭgna (ръгна) mušna (мушна), promuša (промуша)
strike	ucelja (уцеля)
tap (also dab)	potupam (потупам)
throw	xvŭrlja (хвърля)

Appendix B

Tables of the results from the analyses of the English corpus data

VERB	USAGE			
	Syntax		Semantics	
	Active	Passive	Concrete Physical Impact	Extended Meaning
hit	78	22	79	21
smash	79	21	85	15
strike	79	21	57	43
tap	91	9	71	29
dab	100	-	100	-
slap	96	4	95	5
kick	95	5	86	14
stab	60	40	92	8
push	88	12	76	24
pull	93	7	87	13
drag	81	19	82	18
haul	78	22	89	11
throw	95	5	49	51
cast	83	17	43	67
shoot	52	48	92	8
scratch	89	11	93	7
cut	78	22	68	32

Table 1. Results from the English corpus data - usage

VERB	SUBJECT					DIRECT OBJECT						
	Semantics					Syntax		Semantics				
	<i>Initiator</i>	<i>Source</i>	<i>Source extension</i>	<i>Limit</i>	<i>Absorber</i>	<i>Realized</i>	<i>Non-realized</i>	<i>Absorber</i>	<i>Limit</i>	<i>Source extension</i>	<i>Object of Creation</i>	<i>Other</i>
hit	52	22	4	21	1	73	5	2	71	-	1shot	-
smash	60	7	1	-	32	63	16	56	2	3	1	1
strike	49	34	2	14	7	60	21	18	42	-	blow deal hour	-
tap	79	3	8	1	9	74	17	22	39	7	4	2
dab	100	-	-	-	-	57	43	17	37	3	-	-
slap	91	4	-	4	1	91	5	24	62	4	-	1
kick	91	1	3	3	2	77	19	37	30	7	2 hole	-
stab	51	5	3	40	-	47	13	2	41	4	-	-
push	83	3	2	-	12	81	7	76	-	-	-	5
pull	85	8	-	-	7	67	26	66	-	-	-	1
drag	67	13	1	-	19	73	8	69	-	2	-	2
haul	53	25	-	-	22	77	1	77	-	-	-	-
throw	78	20	-	-	5	99	1	97	-	2	-	-
cast	61	22	-	-	17	96	4	96	-	-	-	-
shoot	57	-	1	-	48	42	17	11	29	-	1	-
scratch	81	2	6	-	11	73	16	26	44	-	2	1
cut	67	6	1	-	22	67	11	67	-	-	1	-

Table 2. Results from the English corpus data – Subject and Direct Object

VERB	COMPLEMENTATION																			
	Syntax			Semantics																
				Arguments										Adjuncts						
	AdvP	PP	Sentence	Source	Source extension	Absorber	Limit	Place of Contact	Path				Recipient	Resultative	Manner	Location	Time	Contingency	Quantification	Modality
Origin									End	Length	Orientation									
hit	24	56	11	14	10	-	5	11	2	-	-	1	-	-	15	7	18	3	3	2
smash	4	57	8	4	3	-	19	1	-	2	5	-	-	6	7	-	11	6	1	1
strike	18	56	13	10	5	-	8	3	-	-	-	-	1	1	25	7	12	3	5	2
tap	17	71	5	1	8	6	10	12	4	2	3	2	-	-	18	2	7	-	6	1
dab	17	92	3	-	30	7	46	-	-	-	-	-	-	2	15	2	2	5	3	-
slap	25	72	3	2	5	-	26	23	1	2	1	1	-	1	28	-	3	7	2	-
kick	18	51	9	2	1	-	3	8	3	9	2	14	1	2	13	3	13	2	2	
stab	22	113	5	4	15	-	12	25	-	1	-	1	-	15	12	20	21	-	9	4
push	51	71	12	-	4	-	2	-	8	32	7	32	-	7	20	4	5	9	2	2
pull	39	54	15	-	-	5	1	-	12	16	2	23	-	4	21	12	7	5	-	-
drag	32	81	9	1	-	-	-	2	20	24	14	37	-	3	11	-	13	4	1	2
haul	38	73	12	7	1	-	-	-	27	19	4	25	-	3	12	2	10	7	3	
throw	16	68	1	-	-	-	-	-	8	39	3	13	1	3	7	1	3	5	-	2
cast	16	71	7	2	-	-	-	-	2	46	-	13	2	-	13	3	7	3	1	2
shoot	15	80	6	10	5	-	6	3	-	5	1	5	1	9	10	21	16	7	1	2
scratch	15	39	4	1	7	-	12	1	5	-	-	1		1	12	2	5	6	2	3
cut	16	61	10	2	4	1	-	1	-	3	3	2	2	2	11	2	7	8	1	3

Table 3. Results from the English corpus data - other

Appendix C

Tables of the results from the analyses of the Bulgarian corpus data

VERB	USAGE						
	Syntax					Semantics	
	Active			Passive		Concrete Physical Impact	Extended Meaning
	Default	Reflexive	Absolutive	Se-passive	Periphrastic		
udarja	82	1	14	-	3	87	13
capna	38	1	-	-	5	36	8
ucelja	89	-	-	4	7	92	8
potupam	93	6	-	1	-	100	-
pljasna	85	13	-	-	2	95	5
ritna	95	-	-	-	5	97	3
proboda	81	3	-	1	15	84	16
rūgna	38	1	5	-	-	41	3
mušna	65	33	-	-	2	100	-
butna	98	-	-	-	2	84	16
blūsna	60	-	34	-	6	77	3
drūpna	82	-	18	-	-	75	25
teglja	83	4	4	2	7	30	70
vlača	55	-	42	-	3	88	12
sūborja	81	-	4	-	15	89	11
metna	61	-	35	-	4	92	8
xvūrlja	70	-	25	-	4	58	42
streljam	91	6	-	-	3	96	4
draskam	48	2	-	-	-	47	3
reža	83	-	-	15	2	77	23

Table 1 Results from the Bulgarian corpus data

VERB	SUBJECT					DIRECT OBJECT						
	Semantics					Syntax			Semantics			
	Initiator	Source	Source extension	Limit	Absorber	Realized		Non-realized	Absorber	Limit	Source extension	Other
						Noun Phrase	Clitic (Acc)					
udarja	68	7	8	3	14	33	31	29	3	55	5	1liquid 1hour 1stitch
capna	36	1	2	5	-	13	25	1	2	32	3	2liquid
ucelja	67	3	19	11	-	53	25	11	-	79	-	-
potupam	91	-	2	7	-	51	41	8	-	92	-	-
pljasna	76	5	3	14	2	13	15	57	-	27	1	-
ritna	87	-	-	5	-	44	46	5	81	6	-	-
proboda	59	-	22	19	-	41	39	1	-	80	-	-
rŭgna	38	-	-	-	5	16	18	4	7	27	2	-
mušna	65	-	-	-	35	51	10	4	51	8	2	-
butna	96	-	2	-	2	68	29	1	74	14	1	8 sum
blŭsna	45	11	4	6	34	27	33	-	16	42	1	1 drug
drŭpna	73	8	1	-	18	54	22	6	76	-	-	-
teglja	66	17	-	-	17	54	17	12	71	-	-	32 lotto
vlača	55	-	-	-	45	30	23	2	53	-	-	-
sŭborja	58	21	2	-	19	44	37	-	81	-	-	-
metna	58	2	1	-	39	43	18	-	61	-	-	-
xvŭrlja	60	10	-	-	30	56	14	-	70	-	-	look
streljam	94	-	6	-	-	1	1	-	1	-	-	1 hate
draskam	31	4	-	3	12	9	3	16	11	-	-	1creation
reža	73	5	10	-	12	62	12	9	74	-	-	-

Table 2. Results from the Bulgarian corpus data - Subject and Direct Object

VERB	COMPLEMENTATION																		
	Syntax				Semantics														
					Arguments								Adjuncts						
	Dative Clitic	Adverb	Prep. Phrase	Sentence	Source	Source extension	Limit	Place of Contact	Path				Recipient	Manner	Location	Time	Contingency	Quantification	Modality
Origin									End	Length	Orientation								
udarja	4	27	96	6	2	27	28	31	-	1	-	-	-	21	2	9	4	7	2
capna	2	9	38	6	-	9	-	24	-	2	-	-	2	7	3	3	1	5	3
ucelja	-	20	32	3	1	6	-	14	-	-	-	-	-	9	3	8	1	5	7
potupam	-	10	82	3	-	10	1	65	-	-	-	-	-	13	-	4	4	2	-
pljasna	-	49	107	7	2	49	3	33	-	7	-	-	-	32	-	4	6	10	-
ritna	-	15	76	4	3	15	-	27	1	2	-	7	-	19	6	14	3	-	1
proboda	-	31	84	5	4	31	-	28		1	1	-		15	4	10	5	5	13
rūgna	-	17	47	-	-	17	11	15	-	1	-	1	-	3	-	-	3	1	-
mušna	-	6	95	7	-	6	1	1	3	71	1	3	-	12	-	10	7	-	1
butna	5	8	52	3	-	8	-	2	2	26	5	23	5	8	-	5	2	1	-
blūsna	2	6	95	2	4	6	2	8	1	43	-	12	-	22	13	8	1	-	-
drūpna	-	2	38	13	-	2	-	7	6	-	3	17	-	26	2	5	8	3	2
teglja	6	4	43	7	2	4	-	-	12	2	-	11	6	11	3	14	6	1	3
vlača	-	1	69	7	-	1	-	2	-	2	31	19	-	27	4	12	2	3	6
sūborja	-	5	47	4	1	5	-	-	6	24	-	-	-	7	6	6	2	7	-
metna	2	2	87	6	-	2	-	-	2	69	4	11	-	13	1	13	2	-	1
xvūrlja	3	1	87	11	-	1	-	-	3	56	5	24	-	12	2	9	5	1	2
streljam	-	16	93	3	-	16	41	-	7	-	1	5	-	18	7	9	4	6	3
draskam	-	9	43	3	-	9	21	3	-	-	-	-	-	9	6	1	2	1	1
reža	9	12	45	3	-	12	-	2	-	-	-	-	9	27	3	7	6	1	2

Table 3. Results from the Bulgarian corpus data – other

Appendix D:

Sentence	Usage		Types of Continuations										Conjunction	No Continuation				
	Concrete	Extended Meaning	Direct Object					Complementation										
			Limit	Place of Contact	Absorber	Source Extension	Creation	Limit	Place of Contact	Source Extension	Path	Manner			Location	Time	Reason/Cause	Other
<i>Bob hit</i>	25	5	18	1	7	1	1	5	3	-	3	1	1	1	-	-		
<i>Peter was hitting</i>	28	2	25	-	3	2	-	2	1	-	4	1	2	1	-	-		
<i>The girl was knocking</i>	30	-	-	-	1	-	-	-	-	27	-	-	-	4	-	1		
<i>Frank struck</i>	25	5	24	1	1	1	3	2	1	1	3	1	1	5	1	-	2	
<i>Tom tapped</i>	28	2	7	1	1	5	2	3	-	15	1	4	-	4	-	-	2	
<i>Lucy slapped</i>	30	-	22	3	5	-	-	1	-	-	8	1	1	-	-	1	4	
<i>Bill stabbed</i>	29	1	20	-	-	3	-	8	-	10	4	-	-	1	-	1	2	
<i>Sam kicked</i>	30	-	24	-	6	-	-	-	3	-	-	1	2	-	-	-	2	
<i>Steven dabbed</i>	28	2	4	-	11	2	-	19	1	-	-	-	-	-	-	-	1	
<i>David knocked</i>	29	1	-	-	8	-	-	17	-	-	-	2	-	1	-	2	9	
<i>Terry hammered</i>	30	-	-	-	17	1	-	17	-	-	-	-	-	1	-	2	1	
<i>Harry smashed</i>	29	1	-	-	29	-	-	3	-	-	-	1	1	3	-	1	-	
<i>John scratched</i>	30	-	6	22	2	-	-	-	2	-	2	-	-	2	-	-	1	2
<i>The girl threw</i>	26	4	-	-	25	1	2	-	1	4	7	-	-	1	2	2	2	-
<i>The man cast</i>	21	8	-	-	26	-	1	-	2	6	2	-	-	-	-	-	4	-
<i>Tom shot</i>	22	8	17	1	2	1	-	3	2	1	3	1	1	1	1	2	-	-
<i>Margaret cut</i>	28	1	-	10	18	-	-	-	7	-	-	3	1	1	-	1	-	1

Table 1. Results from the online experiment in English for sentences of the type Subject_[Initiator] Verb _____

Sentence	Usage		Types of Continuations															Conjunction	No Continuation							
	Concrete	Physical Impact	Complementation																							
			Limit	Place of Contact	Absorber	Source	Extension	Creation	Path			Manner	Location	Time	Reason/Cause	Quantification	Result			Other						
Mark hit the door	30	-	-	-	13	-	-	-	-	-	11	-	-	-	-	3	1	1	-	-	-	-	-	-	2	-
Peter struck the wall	30	-	-	-	20	1	-	-	-	-	7	-	-	-	-	-	1	1	-	-	-	-	-	-	1	-
Lucas was knocking on the door	30	-	-	-	1	-	-	-	-	-	8	-	-	-	-	7	5	1	-	-	-	-	-	-	5	-
Jack tapped the table	30	-	-	-	13	-	-	-	-	-	12	-	-	-	-	-	2	2	-	-	-	-	-	-	1	1
Brian stabbed the man	30	-	13	-	7	-	-	-	-	-	1	-	-	-	-	1	1	3	1	-	-	-	3	1	1	-
Harry kicked the boy	30	-	17	-	-	-	-	-	-	-	6	-	-	-	-	3	1	1	-	-	-	-	-	-	-	-
Henry dabbed him	30	-	7	4	7	-	-	-	-	-	3	-	1	-	-	2	1	-	-	-	-	-	-	-	1	1
Mary knocked the chair	30	3	-	-	-	-	-	-	1	1	4	-	1	-	-	-	-	-	-	-	-	-	-	-	1	-
Frank hammered the nail	30	27	-	-	1	-	-	-	1	-	1	-	-	-	-	-	2	-	-	-	-	-	-	-	1	-
Nick smashed the mirror	30	4	-	-	6	-	-	-	-	-	7	-	-	-	-	1	1	-	-	-	-	-	-	-	8	-
Susan threw the ball	30	-	-	-	-	-	2	-	9	13	3	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-
Ron cast the stick	30	-	-	-	-	-	-	-	1	16	1	-	-	-	-	1	1	-	-	-	-	-	-	-	-	-
The man shot the thief	30	-	16	-	1	-	-	-	-	-	2	1	-	1	1	-	1	1	-	-	-	-	-	-	4	1
Cary scratched her leg	30	4	3	-	4	-	-	-	-	-	2	1	4	5	-	-	-	-	-	-	-	-	-	-	2	-
He cut the tree	30	-	-	-	6	-	-	-	-	-	6	-	-	-	-	-	3	-	-	-	-	-	-	-	4	-

Table 2. Results from the online experiment in English for sentences of the type *Subject*_[Initiator] *Verb Object*_____

Sentence	Usage		Types of Continuations													Conjunction	No Continuation									
	Concrete Physical Impact	Extended Meaning	Direct Object				Complementation																			
			Limit	Place of Contact	Absorber	Creation	Limit	Place of Contact	Length	Origin	End	Orientation	Manner	Location	Time			Reason/Cause	Quantification	Result	Other					
<i>The ball hit</i>	30	-	29	-	-	5	-	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
<i>The disaster struck</i>	-	30	4	-	-	-	-	1	-	4	-	17	-	1	-	-	-	-	-	-	-	-	-	-	1	1
<i>The book slapped</i>	30	-	6	1	-	-	-	12	-	2	-	5	-	2	-	10	-	-	-	-	-	-	-	-	-	-
<i>The sea was slapping</i>	30	-	7	-	-	1	-	-	-	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>The rain hammered</i>	30	-	2	-	-	-	-	-	-	3	-	1	-	1	-	5	down	-	-	-	-	-	-	-	-	-
<i>The chair knocked</i>	29	1	2	-	15	-	7	-	1	3	-	2	-	2	-	2	over	-	-	-	-	-	-	-	-	-
<i>The glass smashed</i>	30	-	2	-	-	-	19	-	-	-	-	3	-	1	-	3	-	-	-	-	-	-	-	-	2	-
<i>The machine cast</i>	7	21	-	-	24	5	-	-	7	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>The gun shot</i>	12	-	4	-	4	-	1	-	-	1	-	1	-	1	-	1	off	-	-	-	-	-	-	-	-	-
<i>The car shot</i>	-	30	-	-	1	-	-	1	11	2	-	2	-	2	-	2	-	-	-	-	-	-	-	-	-	-
<i>This surface scratches</i>	28	-	-	-	5	-	-	-	-	17	-	2	-	-	-	-	1	off	-	-	-	-	-	-	1	1
<i>The knife cut</i>	30	-	-	-	14	-	-	4	-	7	-	8	-	4	-	3	-	-	-	-	-	-	-	-	-	-

Table 3. Results from the online experiment in English for sentences of the type **Subject Verb** _____

Appendix E

Sentence	Usage		Types of continuations														No Continuation							
	Concrete Physical Impact	Extended Meaning	Direct Object						Complementation															
			Limit	Place of Contact	Absorber	Source	Extension	Creation	Limit	Place of Contact	Absorber	Source	Extension	Length	Origin	End		Orientation	Manner	Location	Time	Reason/Cause	Other	
<i>Toni udari</i>	26	4	17	-	1	-	3	7	3	-	5	-	2	-	3	-	-	-	-	-	-	-	-	-
<i>Petür udrijaše</i>	30	-	8	-	1	-	-	13	-	-	6	-	1	-	9	-	1	-	1	-	-	-	-	1
<i>Momičeto tropaše</i>	30	-	-	-	-	-	-	14	-	-	18	-	-	-	6	1	1	1	-	-	-	-	-	-
<i>Filip uceli</i>	26	4	26	1	-	-	-	2	2	-	3	-	-	-	1	-	-	-	-	-	-	-	-	4
<i>Angel potupa</i>	30	-	29	-	-	-	-	1	20	-	-	-	-	-	2	-	-	-	-	-	-	-	-	-
<i>Lili pljasna</i>	30	-	5	-	1	-	4	6	3	-	19	-	1	-	3	-	-	-	-	-	-	-	-	-
<i>Bojan probode</i>	30	-	25	5	-	-	-	2	1	-	15	-	-	-	2	-	1	-	-	-	-	-	-	-
<i>Stefan ritna</i>	28	2	26	-	2	-	-	1	1	-	1	-	1	1	1	-	1	-	-	-	-	-	-	-
<i>Jordan bläsna</i>	30	-	17	-	11	-	-	2	1	-	1	1	1	1	5	-	-	-	-	-	-	-	-	-
<i>Todor začuka</i>	29	1	-	-	19	-	-	14	-	-	1	-	-	-	1	-	1	-	-	-	-	-	-	-
<i>Petür sūbori</i>	30	-	-	-	30	-	-	-	-	-	-	-	3	9	-	1	-	-	-	-	-	-	-	-
<i>Momičeto xvürli</i>	25	5	-	-	26	-	-	-	-	-	-	-	-	9	3	1	-	-	-	-	-	-	-	4
<i>Čoveküt metna</i>	28	2	-	-	30	-	-	-	-	-	-	4	10	2	-	-	-	-	-	-	-	-	-	2
<i>Toni streljaše</i>	30	-	-	-	1	-	-	7	-	-	10	-	-	1	14	-	-	-	-	-	-	-	-	-
<i>Ivan odraska</i>	30	-	-	-	30	-	-	-	2	-	3	-	-	-	1	-	-	-	-	-	-	-	-	-
<i>Maria otrjaza</i>	25	5	-	-	18	-	-	-	-	4	-	-	-	-	-	-	-	-	-	-	-	-	-	2

Table 1. Results from the online experiment in Bulgarian for sentences of the type Subject[initiator] Verb ____

Sentence	Usage		Types of continuations														Conjunction	No Continuation
	Concrete	Physical Impact	Complementation															
			Limit	Place of Contact	Source Extension	Path			Manner	Location	Time	Reason/Cause	Quantification	Other				
<i>Stojan udari vratata</i>	30	-				-	18	-							-	-	-	10
<i>Petko uceli stenata</i>	30	-	-	24	-	-	-	2	2	-	-	-	-	-	-	2	-	3
<i>Šašo potupa po masata</i>	30	-	-	14	-	-	-	4	4	-	-	-	-	-	3	1	2	5
<i>Borjana rina muža</i>	30	-	18	-	-	-	-	2	2	1	-	-	1	-	1	-	7	3
<i>Bobi probode mesoto</i>	30	-	1	26	-	-	-	3	3	-	-	-	1	-	1	-	1	-
<i>Straxil promuši čoveka</i>	30	-	6	19	-	-	-	4	4	-	-	-	-	-	-	1	2	-
<i>Tišo začuka piron</i>	30	28	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	1
<i>Marija blūsna stola</i>	30	7	-	4	-	-	4	7	7	-	-	-	2	-	2	-	2	4
<i>Sonja xvūri topkata</i>	30	-	-	1	3	1	8	2	2	-	-	-	-	-	-	-	-	7
<i>Rumen metna pručkata</i>	30	-	-	-	1	-	9	2	2	-	-	-	-	-	-	-	-	2
<i>Čovekūt prostrelja kradeca</i>	30	-	18	4	-	-	-	4	4	-	-	2	-	-	-	1	1	3
<i>Policajāt zastrelja kradeca</i>	30	-	6	-	-	-	-	12	12	-	-	1	1	1	1	3	1	5
<i>Tja odraska kraka si</i>	30	17	-	7	-	-	-	1	1	-	-	2	-	-	-	1	-	2
<i>Toj otrjazā dūrvoto</i>	30	-	-	7	-	-	-	5	5	-	-	-	2	-	2	11	1	4

Table 2. Results from the online experiment in English for sentences of the type **Subject**_[Initiator] **Verb Object**_[Absorber] —

Sentence	Usage		Types of continuations														Conjunction	No Continuation							
	Concrete Physical Impact	Extended Meaning	Direct Object				Complementation												Other						
			Limit	Absorber	Source	Extension	Creation	Limit	Absorber	Place of Contact	Source	Extension	Length	Origin	End	Orientation				Manner	Location	Time	Reason/Cause	Quantification	
<i>Topčeto udari</i>	30	-	26	-	-	4	-	4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>Topkata uceli</i>	30	-	29	-	-	1	-	2	-	-	-	-	-	-	-	1	-	-	-	-	-	-	1	-	-
<i>Knigata pijasna</i>	30	-	4	-	-	2	-	4	-	-	-	-	-	-	18	-	5	-	1	-	-	-	-	-	2
<i>Vühnite pljaskaxa</i>	30	-	9	1	1	11	-	2	-	-	-	-	-	-	-	6	2	-	-	-	-	-	-	2	1
<i>Džždūt čukaše</i>	30	-	-	-	-	28	-	-	-	-	-	-	-	-	-	3	-	-	-	-	-	-	-	1	-
<i>Stolūt se blšsna</i>	30	-	-	-	-	26	-	-	-	-	-	-	-	-	1	4	-	-	-	-	-	-	-	1	-
<i>Butilkata se sūbori</i>	30	-	-	-	-	-	-	-	-	-	-	-	-	-	7	12	6	-	-	-	2	-	-	3	-
<i>Mašinata mjataše</i>	29	1	-	30	-	-	-	-	-	-	-	-	-	-	4	2	1	-	-	-	-	-	-	-	-
<i>Puškata streljaše</i>	30	-	-	-	-	3	2	-	-	-	-	-	-	-	-	-	25	-	-	-	-	-	1	-	-
<i>Kolata se izstrelja</i>	-	30	-	-	-	-	-	-	-	-	-	-	-	6	3	3	11	1	-	-	-	-	-	2	-
<i>Tazi povūrxnost se draska</i>	30	-	-	-	-	-	-	-	-	-	-	-	4	-	-	-	25	-	-	-	-	-	1	-	-
<i>Nožūt režeše</i>	30	-	-	6	-	2	-	-	-	-	-	-	-	-	-	-	21	-	-	-	-	-	-	2	-

Table 3. Results from the online experiment in Bulgarian for sentences of the type *Subject Verb* _____